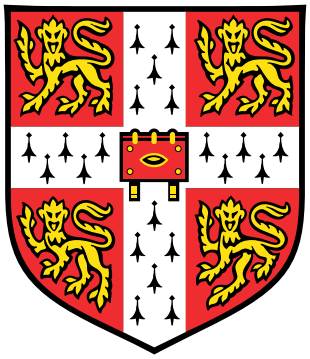
Neutrino interaction vertex reconstruction and particle identification in the MicroBooNE detector

Jack Anthony Weston

Cavendish Laboratory

King's College

April 2021



This thesis is submitted for the degree of Doctor of Philosophy

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Neutrino interaction vertex reconstruction and particle identification in the MicroBooNE detector

Jack Anthony Weston

April 2021

Abstract

This thesis presents the results of a study measuring and improving the quality of neutrino interaction vertex reconstruction and particle identification (PID) in the MicroBooNE detector. The detector comprises a liquid argon time-projection chamber (LArTPC) with a light-collection system, permitting precise tracking of neutrino interaction final states. MicroBooNE's primary physics goal is to resolve the low-energy electron neutrino appearance anomalies observed at MiniBooNE and LSND. The experiment therefore requires high-quality neutrino interaction vertex reconstruction and PID, which together strongly influence event reconstruction quality and energy/momentum estimation. Improvements to the vertex reconstruction are made through the development of powerful new variables and the application of machine learning techniques; these algorithms are now the default used at MicroBooNE and have enabled new studies of neutrino interactions with up to six charged particles in the final state. A robust PID method (FOMA) is developed using a novel analytic approximation to the mode of the $\mathrm{d}E/\mathrm{d}x$ distribution. A deep learning PID method (PidNet) is also proposed, based on convolutional neural networks (CNNs) and a semi-supervised representation learning method. The performance of the two approaches is compared and contrasted with PIDA, the default PID algorithm used at MicroBooNE. This work concludes by assessing the impact of the tools and methods developed in this work on particle energy estimation in MicroBooNE.

Acknowledgements

I would first and foremost like to thank my supervisor, Mark Thomson, for the invaluable guidance and countless insights he provided throughout my time at Cambridge. Special thanks are also owed to Lorena Escudero Sánchez for taking the lead in my supervision in my final years. Lorena went far beyond the call of duty to provide me with academic support and for that I am very grateful. I would also like to thank Leigh Whitehead and Melissa Uchida for their generous help during the final throes of this thesis. John Marshall had a significant hand in my education as a software developer and ignited within me a new passion and attention to detail that will endure for the rest of my life.

During my now decade-long university education, I have been lucky to have been mentored by a number of inspiring physicists and mathematicians. I would like to express particular gratitude to Peter Norreys at the Rutherford Appleton Laboratory for his past and continuing support, and to Tony Weidberg, Andrei Starinets, Devinder Sivia, Yang-Hui He and Georg Viehhauser at Oxford University for their supervision (and endless patience) during my undergraduate years. For making my trip to Fermilab almost enjoyable, I am grateful to my colleagues – and I daresay friends – Andy Smith, Rhiannon Jones and Tom Brooks. I give thanks to my colleagues at the Cavendish, Joris de Vries, Steven Green, Boruo Xu and Stefano Vergani, for making the lab an even more intriguing place to work. For permitting this thesis to compete for my attention these past two years, I am likewise grateful to my partners at Novoic, Emil Fristed and Márton Mészáros, whose 'speed over principle' attitude was alas not reflected in its completion.

Finally, I would like to thank my partner Valerio, my friends and my family for their unwavering support throughout the long process that has been my education. I am indebted to them all.

Contents

List of figures

List of tables

List of abbreviations

ADC	Analogue-to-digital converter
ASIC	Application-specific integrated circuit
AUC	Area under the receiver operating characteristic (ROC) curve
BDT	Boosted decision tree
BNB	Booster Neutrino Beam at Fermi National Accelerator Laboratory (FNAL)
BSM	Beyond the Standard Model (physics)
CC	Charged-current (interaction)
CCDIS	Charged-current deep inelastic scattering
CCQE	Charged-current quasi-elastic (interaction)
CCRES	Charged-current resonant (interaction)
CNN	Convolutional neural network
CP	Charge-parity (symmetry)
CRT	Cosmic-ray tagger (at MicroBooNE)
CSDA	Continuous slowing-down approximation
DAQ	Data acquisition system
DUNE	The Deep Underground Neutrino Experiment

EXT	External trigger (at MicroBooNE)
FNAL	Fermi National Accelerator Laboratory
FOMA	First-order modal approximation
FSI	Final-state interaction
FWHM	Full width at half maximum
GALLEX	Gallium Experiment
GNO	Gallium Neutrino Observatory
LArTPC	Liquid argon time-projection chamber
LHS	Left-hand side
LSND	Liquid Scintillator Neutrino Detector
MAD	Mean absolute deviation
MC	Monte Carlo (simulation)
MIP	Minimum ionizing particle
ModBox	Modified Box (model of recombination)
MPV	Most probable value
NC	Neutral-current (interaction)
NuMI	Neutrinos at the Main Injector facility at Fermi National Accelerator Laboratory (FNAL)
PFO	Particle flow object (in Pandora pattern recognition)
PID	Particle identification
PidNet	Particle identification neural network
PMNS	Pontecorvo–Maki–Nakagawa–Sakata

PMT	Photomultiplier tube
POT	Protons on target
RHS	Right-hand side
RMS	Root mean square
RNN	Recurrent neural network
ROC	Receiver operating characteristic (curve)
SAGE	Soviet–American Gallium Experiment
SCE	Space-charge effect
SDK	Software development kit
SM	The Standard Model (of particle physics)
SNO	Sudbury Neutrino Observatory
SVM	Support vector machine
t-SNE	t-distributed stochastic neighbour embedding
T2K	Tokai-to-Kamioka experiment

Chapter 1

Introduction

Since the existence of the Higgs boson was confirmed at the LHC in 2012, all particles of the Standard Model (SM) have now been observed [1]. However, a number of physical phenomena remain unexplained; for instance, the theory is widely considered to be irreconcilable with general relativity [2] and lacks any viable dark matter candidate [3]. Whilst the SM is thought to be self-consistent and has provided a number of since-confirmed predictions[1] [1, 4–8], it is certainly an incomplete theory.

Of particular relevance to this thesis is the problem posed by neutrinos: the discovery of neutrino oscillations by the Super-Kamiokande and SNO experiments near the turn of the century implies that neutrinos have nonzero mass, contrary to the SM prediction [9, 10]. The Higgs mechanism does not endow left-handed neutrinos with mass if they have no right-handed counterpart. Even if we include Dirac mass terms for neutrinos in the SM Lagrangian, they must be of a size much smaller than for other fermions, at least five orders of magnitude based on current experimental constraints, which similarly has no confirmed theoretical justification. One elegant explanation of both the origin and scale of neutrino masses is the *seesaw mechanism*, which predicts the existence of a heavy, right-handed, sterile[2] partner to the neutrino that can be used to invoke the Higgs mechanism and justify a large mass difference between the left- and right-handed states [11, 12]. Such a particle has yet to be observed but experiments are already

[1] The W and Z bosons, the gluon, the Higgs boson, and the top and charm quarks were all predicted by the Standard Model before they were experimentally observed.

[2] I.e. one that does not interact weakly but does interact gravitationally.

ruling out mass regions [13] while anomalous results at LSND and MiniBooNE are being considered as possible evidence of such a particle [14, 15].

Building and instrumenting detectors capable of accurately reconstructing neutrino interactions is a challenge, but recent developments in neutrino detector technologies are making it possible to identify final state particles with increasing accuracy [16, 17]. Studies of long baseline neutrino oscillations hope to determine the mass hierarchy and probe CP violation in the neutrino sector, while short baseline oscillation analyses will be able to further constrain or rule out sterile neutrino oscillations [18, 19]. The short- and long-baseline neutrino programmes at the Fermi National Accelerator Laboratory (Fermilab) in Illinois make use of liquid argon time-projection chambers (LArTPCs) to address these questions [20]. LArTPCs provide bubble-chamber-like spatial resolution and accurate calorimetry throughout the active detector medium. This allows unprecedentedly high-quality particle tracking, energy estimation and particle identification, including electron-photon discrimination. This thesis had the privilege of analyzing data from one such detector, MicroBooNE, whose main goal is to shed light on the anomalies observed at LSND and MiniBooNE [21]. Improvements presented in this thesis are already contributing towards this effort.

The structure of this work is as follows. **Chapter 2** summarizes the theory of neutrino physics, the theoretical motivation for sterile neutrinos, neutrino detection methods, and a selection of recent and ongoing experiments. In **Chapter 3**, the MicroBooNE experiment is described, including its detector and neutrino sources. **Chapter 4** outlines the Pandora pattern recognition framework, a software package used for event reconstruction at experiments including MicroBooNE. A new algorithm for neutrino interaction vertex selection using a machine learning technique, now the default vertex reconstruction at MicroBooNE, is introduced and validated in **Chapter 5**. **Chapter 6** outlines the theory of energy losses in LArTPCs at MicroBooNE-like energies and derives novel low-energy approximations for $\mathrm{d}E/\mathrm{d}x$. Using one of the low-energy approximations developed in the previous chapter, **Chapter 7** describes and evaluates a new robust PID technique called FOMA. A second, deep learning PID technique is introduced and validated in **Chapter 8**, based on semi-supervised and representation learning paradigms. **Chapter 9** presents an assessment of the impact of the work in this

thesis on particle energy reconstruction at MicroBooNE. Finally, **Chapter 10** provides concluding remarks and discusses future work.

Chapter 2

Neutrino physics

This chapter provides a brief overview of the theory of neutrino physics and current efforts to understand the neutrino. A short description of the Standard Model is given, including notes on the Higgs mechanism, the origin of neutrino mass and the theoretical motivation for sterile neutrinos. The theory of neutrino oscillations is described. The chapter ends with a brief summary of six modern neutrino experiments and a list of open questions in neutrino physics.

2.1 Introduction

Often cited as a historical example of serendipity, Henri Becquerel 'accidentally' discovered radioactivity in 1896 when using uranium to study phosphorescence. He observed that uranium salts were able to blacken a photographic plate, even when wrapped in black paper [22]. A similar observation had been made by Wilhelm Röntgen a few months previously, albeit with an inaccurate explanation [23]. By 1899, Ernest Rutherford had separated these emissions into two categories, alpha and beta, characterized by the differing extents to which they penetrate objects and cause ionization [24].

Beta-decay, however, presented a problem. By 1914, there was strong evidence that the electrons emitted during beta-decay were not monoenergetic as predicted, but rather had a broad, continuous energy spectrum [25]; this apparent two-body problem therefore seemed to violate conservation of energy. It was not until 1930 that Wolfgang Pauli solved the puzzle. In a now-famous letter beginning "*Liebe Radioaktive Damen*

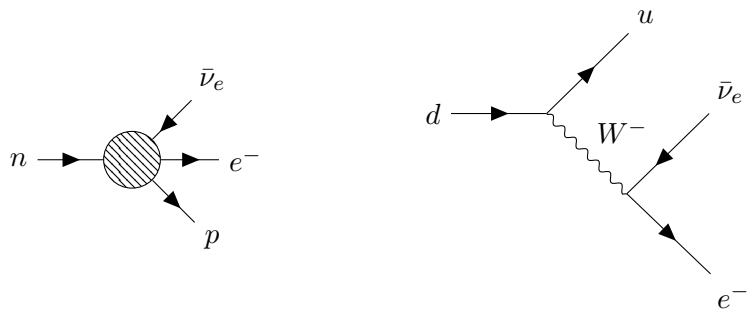


Figure 2.1: Left: a four-fermion point interaction, as used in Fermi's theory of the weak interaction. β^- decay is modelled by a neutron decaying at a single vertex into a proton, electron and electron antineutrino. Right: a modern Feynman diagram for β^- decay via exchange of a W^- boson. β^- decay is modelled by a down-quark decaying to an up-quark and W^- boson, which in turn decays to an electron and an electron antineutrino.

und Herren" ("Dear radioactive ladies and gentlemen"), he theorized the existence of an unseen, neutral, spin-$\frac{1}{2}$ particle that could be used to satisfy conservation of energy, momentum and angular momentum: a particle later christened by Enrico Fermi as the *neutrino* [26]. What was being observed is today known as β^- decay, process that is now well-understood:

$$(Z, N) \to (Z + 1, N - 1) + e^- + \bar{\nu}_e, \tag{2.1}$$

or, on a lower level,

$$n \to p + e^- + \bar{\nu}_e, \tag{2.2}$$

where (Z, N) denotes a nucleus comprising Z protons and N neutrons. This discovery was followed by the discovery of the neutron in 1932 [27] and β^+ decay in 1934 [28].

To explain beta-decay, Fermi proposed his theory of the weak interaction in 1934. In this theory, beta decays are modelled as four-fermion point interactions (see Figure 2.1) with an effective Lagrangian

$$\mathcal{L}_{\text{Fermi}} = \frac{G_F}{\sqrt{2}} \left(\bar{\psi}_p \gamma^\mu \psi_n\right) \left(\bar{\psi}_e \gamma_\mu \psi_\nu\right), \tag{2.3}$$

where G_F is the Fermi constant, ψ_α is the field of particle α and γ_μ are the gamma matrices [29]. This theory allowed the scattering cross-section for neutrinos incident on neutrons to be calculated; this was found by Rudolf Peierls and Hans Bethe in 1934 for

a neutrino of energy E_ν to be

$$\sigma(n + \nu \rightarrow e^- + p) \sim E_\nu \text{ [MeV]} \times 10^{-43} \text{ cm}^2. \tag{2.4}$$

With such small cross-sections, Peierls and Bethe note in [30] that there is "no practically possible way of observing the neutrino" via this process. Luckily, it turns out neutrinos are ubiquitous thanks to solar and atmospheric neutrino spectra, and can be made in vast quantities using nuclear reactors and particle accelerators. Tiny cross-sections have therefore not halted the study of the neutrino: the first experimental observation occurred in 1956 at the Savannah River Site in South Carolina, where Cowan and Reines used a nearby nuclear reactor as a neutrino source for a scintillation detector of cadmium chloride dissolved in water [31]. They observed prompt and delayed scintillation light due to inverse beta decay ($\bar{\nu}_e + p \rightarrow n + e^+$) and subsequent cadmium neutron absorption ($n + {}^{108}\text{Cd} \rightarrow {}^{109\text{m}}\text{Cd} \rightarrow {}^{109}\text{Cd} + \gamma$), respectively, a result that won Reines the Nobel Prize in 1995.

Since then, some key experimental results in neutrino physics have included:

- **Violation of P-symmetry by the weak interaction:** In 1956, Chien-Shiung Wu demonstrated through an experiment involving beta-decay that the weak interaction violated conservation of parity [32], a result that was unexpected at the time.

- **Discovery of the muon-neutrino:** In 1962, Leon M. Lederman, Melvin Schwartz and Jack Steinberger demonstrated the existence of a new flavour of neutrino, the *neutretto* (now known as the muon-neutrino) [33].

- **Discovery of the tau-neutrino:** The tau-neutrino was hypothesized to exist since the discovery of the tau lepton at the Stanford Linear Accelerator Center in 1975 [34]. The first observations of tau-neutrino interactions were announced by the DONUT collaboration in 2000 [35].

- **Discovery of weak neutral currents:** Weak neutral currents allow elastic scattering of neutrinos in matter via exchange of the neutral Z boson. Abdus Salam, Sheldon Glashow and Steven Weinberg predicted neutral currents in the

mid-1960s [36–38] and their existence in neutrino interactions was confirmed in 1973 by the Gargamelle bubble chamber at CERN [4, 5].

- **Discovery of the W and Z bosons:** Whilst other experiments had provided evidence for the existence of the W and Z bosons, they were not directly detected until experiments led by Carlo Rubbia and Simon van der Meer provided unambiguous evidence in 1983 [39, 40].

- **Observation of neutrino oscillations:** A significant ($\sim 2/3$) deficit of electron neutrinos in the solar neutrino spectrum was first observed by the Homestake experiment around 1970 [41]. Clear evidence that this was due to neutrino flavour oscillations, implying a nonzero neutrino mass, was provided by the Sudbury Neutrino Observatory in 2001 [42]. Similar effects were observed for atmospheric neutrinos by Super-Kamiokande in 1998 [10]. Multiple experiments have since observed flavour oscillations in solar, atmospheric, reactor and beam neutrino spectra and measured oscillation parameters.

2.2 The Standard Model

The Standard Model (SM) of particle physics is a non-abelian gauge theory that describes three of the four fundamental forces of nature: the electromagnetic, weak and strong interactions. Its gauge group is

$$G_{\mathrm{SM}} = \mathrm{SU}(3) \times \mathrm{SU}(2)_L \times \mathrm{U}(1), \tag{2.5}$$

where the SU(3) symmetry gives rise to QCD, and the $\mathrm{SU}(2)_L \times \mathrm{U}(1)$ symmetry to electroweak interactions. The model contains 24 fermion fields: six quark flavours, each with three distinct colours, plus six leptons flavours. These are summarized in Table 2.1. A fermion field ψ obeys the Dirac equation

$$\left(i\partial\!\!\!/ - m\right)\psi = 0, \tag{2.6}$$

where $\partial\!\!\!/ := \gamma^\mu \partial_\mu$, γ_μ are the gamma matrices and $\partial_\mu := \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ is the four-vector of partial derivatives. A gauge transformation in the SM is represented by a

Table 2.1: The fermions of the Standard Model, along with their mass and electric charge Q as a multiple of e. All have spin $\frac{1}{2}$ and each quark represents three fields (with colour charge red, green or blue) due to QCD's SU(3) gauge group. The quoted top mass corresponds to direct measurements only. Data from [13].

		Mass (MeV)	Q
quarks	u	$2.16^{+0.49}_{-0.26}$	$+\frac{2}{3}$
	d	$4.67^{+0.48}_{-0.17}$	$-\frac{1}{3}$
	s	93^{+11}_{-5}	$-\frac{1}{3}$
	c	$(1.27 \pm 0.02) \times 10^3$	$+\frac{2}{3}$
	b	$(4.18^{+0.03}_{-0.02}) \times 10^3$	$-\frac{1}{3}$
	t	$(172.76 \pm 0.30) \times 10^3$	$+\frac{2}{3}$
leptons	e^-	$0.5109989461 \pm 0.0000000031$	-1
	ν_e	$< 1.1 \times 10^{-6}$ (90% CL; assuming $m_{\bar{\nu}_e} = m_{\nu_e}$)	0
	μ^-	$105.6583745 \pm 0.0000024$	-1
	ν_μ	< 0.19 (90% CL)	0
	τ^-	$1,776.86 \pm 0.12$	-1
	ν_τ	< 18.2 (95% CL)	0

unitary operator U that transforms a fermion field ψ via

$$\psi \to \psi' := U\psi. \tag{2.7}$$

We may split a fermion field ψ into its left- and right-handed chiral components such that $\psi = \psi_L + \psi_R$, where

$$\psi_L := \frac{1}{2}\left(1 - \gamma_5\right)\psi, \tag{2.8}$$

$$\psi_R := \frac{1}{2}\left(1 + \gamma_5\right)\psi. \tag{2.9}$$

Due to the $\mathrm{SU}(2)_L$ symmetry of the SM, these two components transform differently.

To construct a theory that satisfies the required symmetries given by Equation 2.5, we define a Lagrangian density that is invariant under the SU(3), $\mathrm{SU}(2)_L$ and U(1) gauge transformations. This can be done by modifying the operator in Equation 2.6 to include a number of gauge fields. We require one gauge field for each generator of the symmetry group. A SU(n) group requires $n^2 - 1$ generators; a U(n) group requires

n^2 generators. We therefore introduce $(3^2-1)+(2^2-1)+1^2 = 12$ gauge fields via definition of a covariant derivative

$$\partial^\mu \to D^\mu := \partial^\mu + ig_1 Y B^\mu + ig_2 \boldsymbol{T}\cdot\boldsymbol{W}^\mu + ig_3 \boldsymbol{X}\cdot\boldsymbol{G}^\mu, \tag{2.10}$$

where

- B^μ is the field of the U(1) group, part of the electroweak interaction;
- $W_k^\mu, k = 1,2,3$ are the three fields of the SU(2) group, part of the electroweak interaction;
- $G_k^\mu, k = 1,2,\ldots,8$ are eight fields of the SU(3) group, corresponding to the eight gluons;
- $\boldsymbol{T}$ is the SU(2) group generator, usually chosen as $T_k = \frac{1}{2}\tau_k$, where τ_k are the Pauli matrices;
- $\boldsymbol{X}$ is the SU(3) group generator, usually chosen as $X_k = \frac{1}{2}\lambda_k$, where λ_k are the Gell-Mann matrices;
- Y is the *weak hypercharge* of the particle, (see below); and,
- g_1, g_2 and g_3 are coupling constants.

The weak hypercharge of a particle is a quantum number that represents the strength with which it couples to electroweak interactions. It is related to the electric charge via $Y = Q - I_3^W$, where I_3^W is the third component of the *weak isospin*. For right-chiral neutrino fields, $Q = I_3^W = Y = 0$, such that right-handed neutrinos are entirely decoupled from the SM.

Due to spontaneous symmetry breaking[1], the W_3^μ and B^μ fields mix to create the Z boson (with field Z_μ) and the photon (with field A_μ):

$$\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B^\mu \\ W_3^\mu \end{pmatrix}, \tag{2.11}$$

[1] *Spontaneous symmetry breaking* refers to the apparent loss of a symmetry of a system in a particular state. The symmetries of a system present in its Lagrangian may be clear at higher energies whilst particular solutions at lower energies appear asymmetric.

Table 2.2: The gauge bosons of the Standard Model, along with their spin, mass and charge Q as a multiple of e. Exact values in the table are theoretical. Data from [13].

	Gauge boson	Mass (GeV)	Q	Spin
electromagnetic	γ	$< 1 \times 10^{-18}$	$< 1 \times 10^{-35}$	1
weak	$W^{\pm}$	80.379 ± 0.012	± 1	1
	Z	91.1876 ± 0.0021	0	1
strong	g ($\times 8$)	0	0	1
Higgs	H	125.10 ± 0.14	0	0

where

$$\cos\theta_W = \frac{g_2}{\sqrt{g_1^2 + g_2^2}}, \tag{2.12}$$

$$\sin\theta_W = \frac{g_1}{\sqrt{g_1^2 + g_2^2}}, \tag{2.13}$$

and θ_W is known as the *Weinberg angle*. The coupling constants must further satisfy

$$g_1 \cos\theta_W = g_2 \sin\theta_W = e, \tag{2.14}$$

where e is the elementary charge. The W_1 and W_2 bosons similarly mix to create the W^+ and W^- bosons:

$$W_\mu^{\pm} = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp i W_\mu^2 \right). \tag{2.15}$$

The twelve gauge bosons (photon/γ, $W^{\pm}$, Z^0 plus the eight gluons) are summarized in Table 2.2.

Under a gauge transformation U, a gauge field K_μ transforms via

$$K_\mu \to K'_\mu := U K_\mu U^\dagger + \frac{i}{g} \left(\partial_\mu U \right) U^\dagger, \tag{2.16}$$

where K_μ may be any of B_μ, $\boldsymbol{W}_\mu$ or $\boldsymbol{G}_\mu$ and g is the relevant coupling constant. Since these are spin-1 fields, each of B_μ, $\boldsymbol{W}_\mu$ and $\boldsymbol{G}_\mu$ contributes a term to the overall Lagrangian of the form

$$\mathcal{L} = -\frac{1}{4} F_i^{\mu\nu} F_{\mu\nu i} + \frac{1}{2} m_K^2 K_{i\mu} K_i^\mu, \tag{2.17}$$

where

$$F_i^{\mu\nu} = \partial^\mu K_i^\nu - \partial^\nu K_i^\mu - g f_{ijk} K_k^\mu K_k^\nu, \tag{2.18}$$

m_K is the mass of the gauge boson, K_i^μ is the ith gauge field of the group, and f_{ijk} are the structure constants of the group. The f_{ijk} are defined via commutation relations between the generators of a given group:

$$[T_i, T_j] = i f_{ijk} T_k. \tag{2.19}$$

For U(1), the structure constant is 0. SU(2) and SU(3) have nonzero structure constants; for SU(2), the structure constants are given by the Levi-Civita tensor ϵ_{ijk} [43]. Nonzero structure constants allow self-interactions, which are possible for gluons, $W^\pm$ bosons and Z bosons but not photons.

2.2.1 The Higgs mechanism

We introduce mass terms into the SM Lagrangian via the Higgs mechanism: the addition of a universal field (the Higgs field) that undergoes spontaneous symmetry breaking, allowing electroweak gauge bosons to interact with it and gain mass. We require that this happen in such a way as to produce three massive gauge bosons ($W^\pm$, Z) and one massless gauge boson (γ). Consider a field H that contributes a Lagrangian density

$$\mathcal{L}_H := \left(D'_\mu H\right)^\dagger D'^\mu H - V(H), \tag{2.20}$$

where the potential $V(H)$ is defined as

$$V(H) := -\mu^2 H^\dagger H + \lambda \left(H^\dagger H\right)^2, \tag{2.21}$$

D'^μ is the electroweak part of the covariant derivative defined in Equation 2.10, *viz.*

$$D'^\mu := \partial^\mu + i g_1 Y B^\mu + i g_2 \boldsymbol{T} \cdot \boldsymbol{W}^\mu, \tag{2.22}$$

and $\lambda > 0$ and $\mu^2 > 0$ are constants. We can 'break' the symmetry of the Higgs Lagrangian by expanding its field about a nonzero vacuum expectation value. Doing so generates two gauge boson mass terms:

$$\frac{(g_1 v)^2}{4} W_\mu^+ W^{-\mu} + \frac{\left(g_1^2 + g_2^2\right) v^2}{8} Z_\mu Z^\mu, \tag{2.23}$$

which lead to electroweak boson masses

$$m_W = \frac{g v}{2}, \tag{2.24}$$

$$m_Z = \frac{v\sqrt{g_1^2 + g_2^2}}{2} = \frac{m_W}{\cos\theta_W}, \tag{2.25}$$

$$m_A = 0, \tag{2.26}$$

i.e. three massive bosons and one massless boson, as required. The prediction that $\rho := m_W^2/m_Z^2 \cos^2\theta_W$ should equal 1 has been experimentally tested; ρ has been measured to be 1.00040 ± 0.00024.

Fermions similarly gain mass through the Higgs mechanism. The direct inclusion of a Dirac mass term like $m\bar{\psi}\psi$ is not possible as it violates gauge invariance; however, it is possible to include gauge-invariant terms that, after spontaneous symmetry breaking, result in mass terms. One possibility is a *Yukawa coupling* between a fermion field ψ and the Higgs field with some coupling constant G_Y, resulting in a Lagrangian density like

$$\mathcal{L}_{\text{fermion}} = \bar{\psi}\gamma^\mu D_\mu \psi + G_Y \bar{\psi} H \psi. \tag{2.27}$$

The mass generation then follows the same process as for the gauge bosons. The mechanism has been successful in explaining the masses of fundamental particles and won Peter Higgs and François Englert the Nobel Prize in 2013, following the discovery of the theory's predicted Higgs boson at the Large Hadron Collider at CERN in the same year.

2.2.2 Neutrino mass

Before the observation of neutrino oscillations (see Section 2.3), neutrinos in the SM were assumed to be massless. Neutrinos, being fermions, could be expected to gain mass through a Yukawa coupling like $G_Y \bar{\psi} H \psi$. After spontaneous symmetry breaking, we would then expect to see a Dirac mass term:

$$\begin{aligned}
\mathcal{L} &= m\bar{\psi}\psi \\
&= m\overline{(\psi_L + \psi_R)}\,(\psi_L + \psi_R) \\
&= m\left(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L\right),
\end{aligned}$$

where we have used the result $\bar{\psi}_L\psi_L = \bar{\psi}_R\psi_R = 0$. As mentioned before, right-handed neutrinos do not otherwise appear in the SM; if right-handed neutrinos do not exist, this mass term vanishes and SM neutrinos are rendered massless. One way around this

is to assume that neutrinos are *Majorana particles* (i.e. their own antiparticle). In this case, the right-handed component becomes $\psi_L^C := C\bar{\psi}_L^T$, where C is the charge conjugation operator, and the mass term does not vanish. However, this would mean that the electroweak interaction violates conservation of lepton number. We could alternatively assume the existence of a right-handed *sterile neutrino*, i.e. one that does not interact weakly but may interact gravitationally. This idea is theoretically well-motivated and there is some experimental evidence for it, notably at MiniBooNE and LSND (see Chapter 3).

Allowing for the possibility of right-handed neutrinos, and ignoring the issues above for now, we can construct four different mass terms: a left-handed Majorana mass term with mass m_L

$$\mathcal{L}_{\mathrm{L}}^M := \frac{1}{2} m_L \bar{\psi}_L^C \psi_L + \mathrm{h.c.}, \tag{2.28}$$

a right-handed Majorana mass term with mass m_R

$$\mathcal{L}_{\mathrm{R}}^M := \frac{1}{2} m_R \bar{\psi}_R^C \psi_R + \mathrm{h.c.}, \tag{2.29}$$

a Dirac mass term with mass m_D

$$\mathcal{L}_{\mathrm{L}}^D := m_D \bar{\psi}_R \psi_L + \mathrm{h.c.}, \tag{2.30}$$

and a Dirac mass term made from the charge-conjugate fields (corresponding to the same mass m_D)

$$\mathcal{L}_{\mathrm{R}}^D := m_D \bar{\psi}_R^C \psi_L^C + \mathrm{h.c.}, \tag{2.31}$$

where 'h.c.' denotes the Hermitian conjugate. So the most general expression we can write down for the mass term is

$$2\mathcal{L}_{\mathrm{mass}} = L_L^D + L_R^D + L_L^M + L_R^M + \mathrm{h.c.}, \tag{2.32}$$

which can be written as a matrix equation

$$\mathcal{L}_{\mathrm{mass}} \sim \begin{pmatrix} \bar{\psi}_L^C & \bar{\psi}_R \end{pmatrix} \begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R^C \end{pmatrix} + \mathrm{h.c.}. \tag{2.33}$$

Clearly these are not mass eigenstates, as evidenced by the off-diagonal terms in the mass matrix M. To find the mass eigenstates, we look for a unitary matrix U that transforms

these chiral eigenstates into left-handed field components with definite mass:

$$\begin{pmatrix} \psi_L \\ \psi_R^C \end{pmatrix} = U \begin{pmatrix} \psi_{1,L} \\ \psi_{2,L} \end{pmatrix}, \tag{2.34}$$

and

$$M' := U^\dagger M U. \tag{2.35}$$

U is guaranteed to exist. The only 2×2 unitary matrix is the rotation matrix, so we have

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{2.36}$$

therefore

$$\psi_L = \cos\theta \psi_{1,L} - \sin\theta \psi_{2,L}, \tag{2.37}$$

$$\psi_L = \sin\theta \psi_{1,L} + \cos\theta \psi_{2,L}. \tag{2.38}$$

We choose the rotation angle θ such that U diagonalizes M; from Equation 2.35,

$$M' = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}. \tag{2.39}$$

The result is

$$m_{1,2} = \frac{1}{2}\left[(m_L + m_R) \pm \sqrt{(m_L - m_R)^2 + 4m_D^2}\right]. \tag{2.40}$$

Different values for m_L, m_R and m_D will yield different physical masses m_1 and m_2 but the case $m_L = 0$, $m_R \gg m_D$ is particularly interesting.[2] In this case,

$$m_1 = \frac{m_D^2}{m_R}, \tag{2.41}$$

$$m_2 = m_R\left(1 + \frac{m_D^2}{m_R^2}\right) \simeq m_R, \tag{2.42}$$

i.e. m_2 becomes very large compared with m_1. In terms of the mass eigenstates, we then have

$$\psi_1 \sim \left(\psi_L + \psi_L^C\right) - \frac{m_D}{m_R^2}\left(\psi_R + \psi_R^C\right), \tag{2.43}$$

$$\psi_2 \sim \left(\psi_R + \psi_R^C\right) + \frac{m_D}{m_R^2}\left(\psi_L + \psi_L^C\right), \tag{2.44}$$

[2]Note that the SM explicitly requires $m_L = 0$.

such that the particle ψ_1 is mostly the familiar left-handed neutrino and the particle ψ_2 is mostly its much heavier, sterile, right-handed partner. This class of theories are known as *seesaw mechanisms* and can explain why the neutrino masses are substantially smaller than those of the charged leptons [44]. This mechanism assumes that the neutrino is a Majorana particle and there exists a yet-unobserved heavy, sterile neutrino.

For a more in-depth review of the SM, including the Higgs mechanism and neutrino mass, see [45] or [46].

2.3 Neutrino oscillations

This section outlines some standard theory results in the context of a modern understanding of neutrino oscillation phenomena. For a deeper review of the theory, see [47]; for a comprehensive review of the current experimental status of neutrino oscillations, see [13].

Thanks to neutrino experiments such as Daya Bay and T2K (see Section 2.6), it has been understood since about 2011 that all three known neutrino types undergo flavour oscillation, implying that all three types have nonzero mass. This means that there exists a set of mass eigenstates[3] ν_1, ν_2, ν_3 that may be a mixture of the flavour eigenstates ν_e, ν_μ, ν_τ. The mass eigenstates diagonalize the free Hamiltonian of the neutrino field and therefore describe a neutrino propagating in a vacuum.

This is made mathematically concrete by defining a unitary transformation U that relates the flavour eigenstates ν_α to the mass eigenstates ν_i via

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{2.45}$$

where U in this case takes the form of a 3×3 matrix assuming a three-neutrino model. Since the mass eigenstates are eigenvectors of the free Hamiltonian, a neutrino mass eigenstate ν_i with position four-vector $\boldsymbol{x}$ and four-momentum $\boldsymbol{p}_i$ evolves according to

$$\nu_i(\boldsymbol{x}) = e^{-\boldsymbol{p}_i \cdot \boldsymbol{x}} \nu_i, \tag{2.46}$$

[3] These objects are spinors representing neutrino fields, like ψ in the previous section. In the context of neutrino oscillations, bra-ket notation is often used: $\nu \to |\nu\rangle$ and $\nu^\dagger \to \langle\nu|$ etc.

and the flavour eigenstates therefore evolve as

$$\begin{pmatrix} \nu_e(\boldsymbol{x}) \\ \nu_\mu(\boldsymbol{x}) \\ \nu_\tau(\boldsymbol{x}) \end{pmatrix} = e^{-\boldsymbol{p}_i \cdot \boldsymbol{x}} U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{2.47}$$

which we can write compactly for a single flavour eigenstate ν_α using Einstein notation:

$$\nu_\alpha = e^{-\boldsymbol{p}_i \cdot \boldsymbol{x}} U_{\alpha i} \nu^i, \tag{2.48}$$

noting the implicit sum over i. To get the evolution of a flavour eigenstate, we then use the unitarity property of U to invert Equation 2.45:

$$\nu_\alpha(\boldsymbol{x}) = e^{-\boldsymbol{p}_i \cdot \boldsymbol{x}} U_\alpha{}^i U^*_{\beta i} \nu^\beta. \tag{2.49}$$

This mixing of flavour eigenstates leads to neutrino oscillation phenomena: if neutrinos have mass and the off-diagonal elements of U are nonzero, then we expect there to be a nonzero probability that a neutrino observed as a particular flavour at time t is observed as a different flavour at a later time t'. To make this more concrete, consider a neutrino propagating with $E \gg m_i$:

$$p_i = \sqrt{E^2 - m_i^2} = E - \frac{m_i^2}{2E} + \mathcal{O}\left(\frac{m^4}{E^3}\right). \tag{2.50}$$

If we neglect terms $\mathcal{O}\left(m^4/E^3\right)$ and higher and assume that all mass eigenstates have the same energy E then, at a distance L, the phase in Equation 2.49 becomes

$$\boldsymbol{p}_i \cdot \boldsymbol{x} = Et - p_i L \simeq \frac{m_i^2 L}{2E}, \tag{2.51}$$

where we are using natural units and have assumed the relativistic limit such that $L \simeq t$. The same result can be derived more rigorously, see e.g. [48]. Equation 2.49 therefore becomes

$$\nu_\alpha(\boldsymbol{x}) = e^{-m_i^2 L/2E} U_\alpha{}^i U^*_{\beta i} \nu^\beta. \tag{2.52}$$

The dependence of the phase on m_i means the probability of observing each flavour is not necessarily constant if the neutrinos have distinct masses. We can calculate this probability directly using

$$P\left(\nu_\alpha \to \nu_\beta\right) = \left| \nu_\beta^\dagger \nu_\alpha(\boldsymbol{x}) \right|^2. \tag{2.53}$$

From Equation 2.52,

$$P\left(\nu_\alpha \to \nu_\beta\right) = \delta_{\alpha\beta} - 4\sum_{i>j} \mathrm{Re}\left(U^*_{\alpha i}U_{\beta i}U_{\alpha j}U^*_{\beta j}\right)\sin^2\left(\frac{\Delta m^2_{ij}L}{4E}\right) + \\ + 2\sum_{i>j} \mathrm{Im}\left(U^*_{\alpha i}U_{\beta i}U_{\alpha j}U^*_{\beta j}\right)\sin\left(\frac{\Delta m^2_{ij}L}{2E}\right), \tag{2.54}$$

where $\Delta m^2_{ij} := m^2_i - m^2_j$ and $\delta_{\alpha\beta}$ is the Kronecker delta. The periodicity in L/E is now apparent.

The magnitude of the mixing depends on the entries in the U matrix, which is known as the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix in the neutrino sector [49]. Without loss of generality, we may write U as the product of three matrices that satisfy the unitarity condition by construction:

$$\begin{aligned} U &:= \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned} \tag{2.55}$$

where $c_{ij} := \cos\theta_{ij}$ and $s_{ij} := \sin\theta_{ij}$. The mixing therefore can be expressed as the three mixing angles (θ_{12}, θ_{23} and θ_{13}) and the *Dirac phase* δ. The size of δ controls the extent to which neutrinos violate CP-symmetry. If neutrinos are indeed Majorana particles (see Section 2.2.2), then two more phases must be added to the PMNS matrix. These phases do not affect neutrino oscillations so can be omitted here.

Calculations can be simplified by considering two sets of two-flavour oscillations ($\nu_e \leftrightarrow \nu_\mu/\nu_\tau$ and $\nu_\mu \leftrightarrow \nu_\tau$) separately. This is justified by the experimental observation that θ_{12} and θ_{23} are large, and θ_{13} is small. $\nu_e \leftrightarrow \nu_\mu/\nu_\tau$ oscillations correspond to the first matrix in the decomposed PMNS matrix equation (Equation 2.55) and $\nu_\mu \leftrightarrow \nu_\tau$ to the third matrix. There is experimental evidence that $\left|\Delta m^2_{32}\right| \gg \left|\Delta m^2_{21}\right|$ (though the ordering of the hierarchy is yet to be determined), meaning that these oscillations also occur on different length scales. Over a long range (i.e. $L \sim E/\Delta m^2_{21}$), the oscillation

	Mass ordering	Value
$\sin^2\theta_{12}$	-	$0.307^{+0.013}_{-0.012}$
$\sin^2\theta_{23}$	normal	0.545 ± 0.021
$\sin^2\theta_{23}$	inverted	0.547 ± 0.021
$\sin^2\theta_{13}$	-	$(2.18 \pm 0.07) \times 10^{-2}$
Δm^2_{21}	-	$(7.53 \pm 0.18) \times 10^{-5}$ eV^2
Δm^2_{32}	normal	$(2.453 \pm 0.034) \times 10^{-3}$ eV^2
Δm^2_{32}	inverted	$(-2.546^{+0.034}_{-0.040}) \times 10^{-3}$ eV^2
δ/π	-	1.36 ± 0.17

Table 2.3: Current best estimates of neutrino mixing parameters. 'Normal' mass ordering refers to the ordering $m_1 < m_2 < m_3$; 'inverted' refers to $m_3 < m_1 < m_2$. Data from [13].

probability can be approximated as

$$P(\nu_e \to \nu_\mu/\nu_\tau) \simeq \sin^2(2\theta_{12}) \sin^2\left(\frac{\Delta m^2_{21} L}{4E}\right), \tag{2.56}$$

which can be used to explain ν_e disappearance in solar neutrino spectra. Over a short range (i.e. $L \sim E/\Delta m^2_{32}$),

$$P(\nu_\mu \to \nu_\tau) \simeq \sin^2(2\theta_{23}) \sin^2\left(\frac{\Delta m^2_{32} L}{4E}\right), \tag{2.57}$$

which can be used to explain ν_μ disappearance in atmospheric neutrino spectra. Given a particular mass splitting Δm^2, the apparent disappearance is maximized when $L/E \sim \pi/\Delta m^2$.

A summary of current oscillation parameters is provided in Table 2.3. Measurement of oscillation phenomena is of great interest for understanding neutrinos. Experimental evidence for neutrino oscillations led to a modification of the SM to account for nonzero neutrino masses, providing a theoretical motivation for sterile neutrinos (see Section 2.2.2). Further, CP violation in the neutrino sector could account for a significant part of baryogenesis in the early universe [50, 51]. Alternative or complementary explanations of neutrino disappearance phenomena include quantum decoherence [52, 53] and neutrino decay [54, 55].

2.4 Neutrino sources

Neutrinos can be studied using natural sources, such as solar, atmospheric and supernova [58] spectra, or using manmade sources, such as from nuclear reactors and particle accelerators. Other neutrino sources include the cosmic neutrino background (a relic of the Big Bang) [59] and *geoneutrinos* from radionuclide decay occurring naturally in the Earth [60]. A comparison of the spectra of different neutrino sources is given in Figure 2.2.

Most solar neutrinos come from the initial deuterium production process $p+p \rightarrow {}^{2}\mathrm{H}+ e^{+} + \nu_e$. These neutrinos are low energy ($< 0.42\,\mathrm{MeV}$) and solar neutrino experiments are usually not sensitive to these, but are sensitive to higher-energy neutrinos produced by other fusion processes. A review of solar neutrinos can be found in [61]. Atmospheric neutrinos come from high-energy cosmic rays incident on the Earth's atmosphere, which create hadronic showers whose decay products include neutrinos. Cosmic rays mostly comprise protons, helium nuclei and electrons [62], though those arriving at the surface of the Earth are primarily muons. This process is illustrated in Figure 2.3.

Fission reactors, which break down heavy isotopes like ${}^{235}\mathrm{U}$ to produce energy, induce chains of beta-decays through radioactive decay that generate copious numbers of electron antineutrinos. These processes are well-understood and can be used to create an intense neutrino flux with a well-defined spectrum [63]. Particle accelerators can also be used to create neutrino beams with well-defined spectra. To facilitate study of the beam composition and to measure neutrino oscillations, such experiments typically have both near and far detectors, such that the near detector can be used to characterize the beam. To create a neutrino beam, hadronic showers can be induced through high-energy collisions of protons with a target typically made of graphite, and the neutrinos generated by decays of the pions and kaons contained in the shower can be efficiently isolated from the charged particles [64] to create a very pure beam. A schematic of the NuMI beamline at Fermilab is shown in Figure 2.4 and described in more detail in Section 3.2.

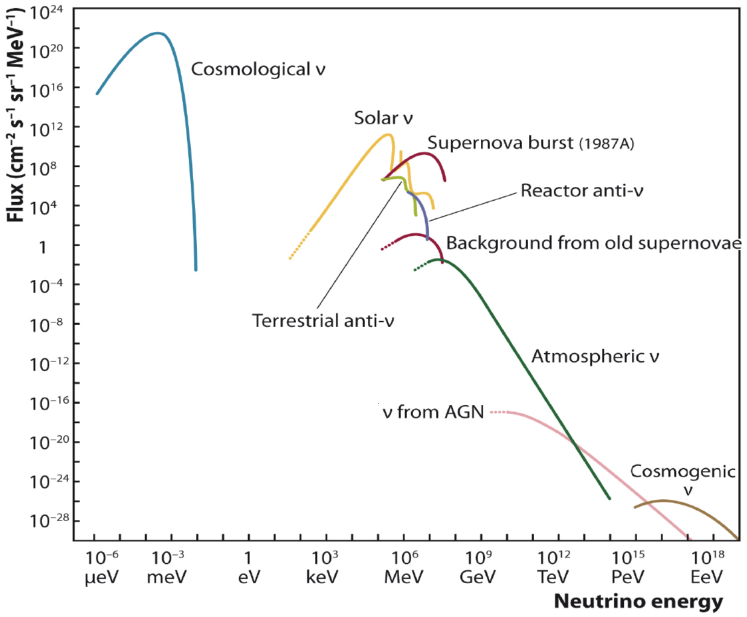


Figure 2.2: A comparison of the neutrino energy spectra provided by a range of different natural/manmade sources. Figure from [56].

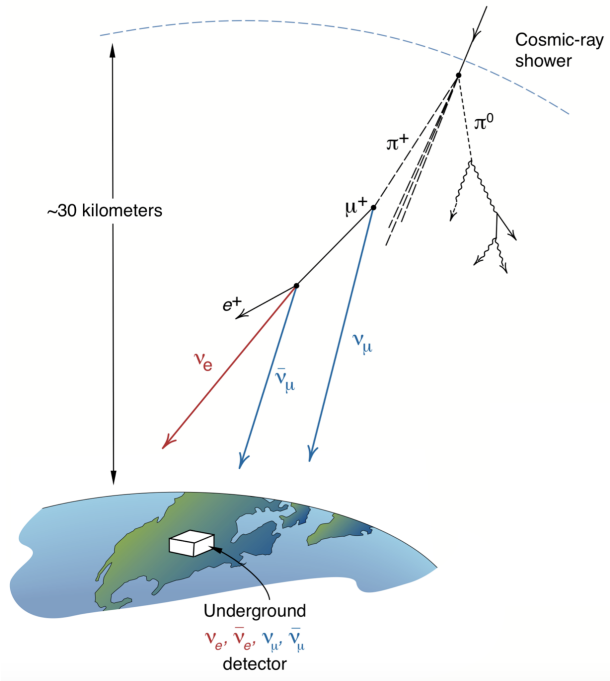


Figure 2.3: Illustration of one process by which neutrinos are generated via interactions of cosmic rays with the Earth's atmosphere. Figure from [57].

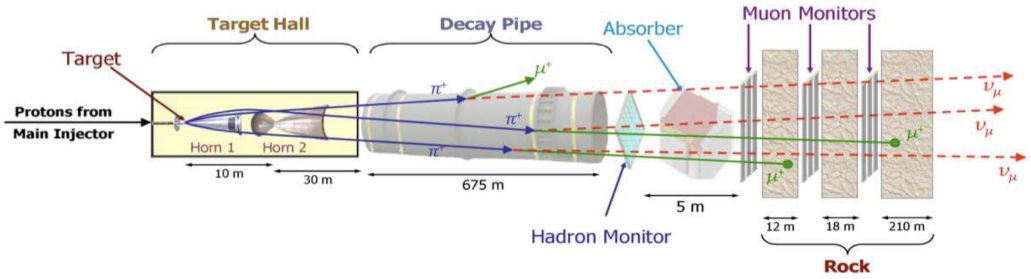


Figure 2.4: A schematic of the NuMI beamline at Fermilab, which provides neutrinos for MINOS, MiniBooNE and ArgoNeuT experiments. Figure from [64].

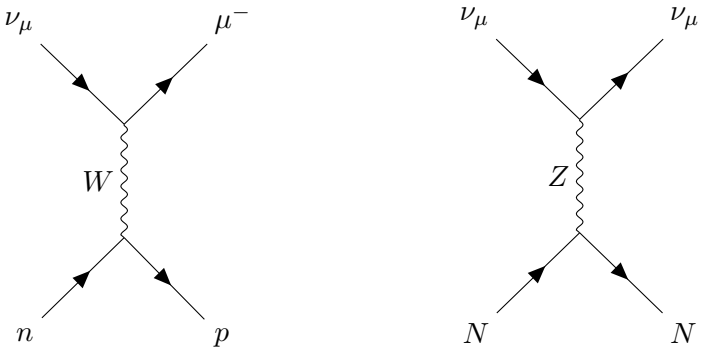


Figure 2.5: Left: a typical charged-current (CC) neutrino interaction, mediated by exchange of a W boson. Right: a typical neutral-current (NC) neutrino interaction, mediated by exchange of a Z boson. N here represents an atomic nucleus.

2.5 Detecting neutrinos

Neutrinos only couple to gravity and the weak interaction. Due to small interaction cross-sections, large detector media and intense fluxes are required to study neutrinos. Since they have no electric charge, they are usually detected indirectly – via the products of their weak interactions with electrons and nuclei. Such interactions can either be *charged current* (mediated by exchange of a W boson) or *neutral current* (mediated by exchange of a Z boson), as shown in Figure 2.5. There are a number of possible modes of interaction with different topological and calorimetric signatures, three of which are shown in Figure 2.6.

Intense neutrino sources used for experimentation include nuclear reactors, the Sun, and neutrino beams created using particle accelerators. As discussed in Section 2.3, the source-to-detector distance also affects observations. In the case of beamline experiments

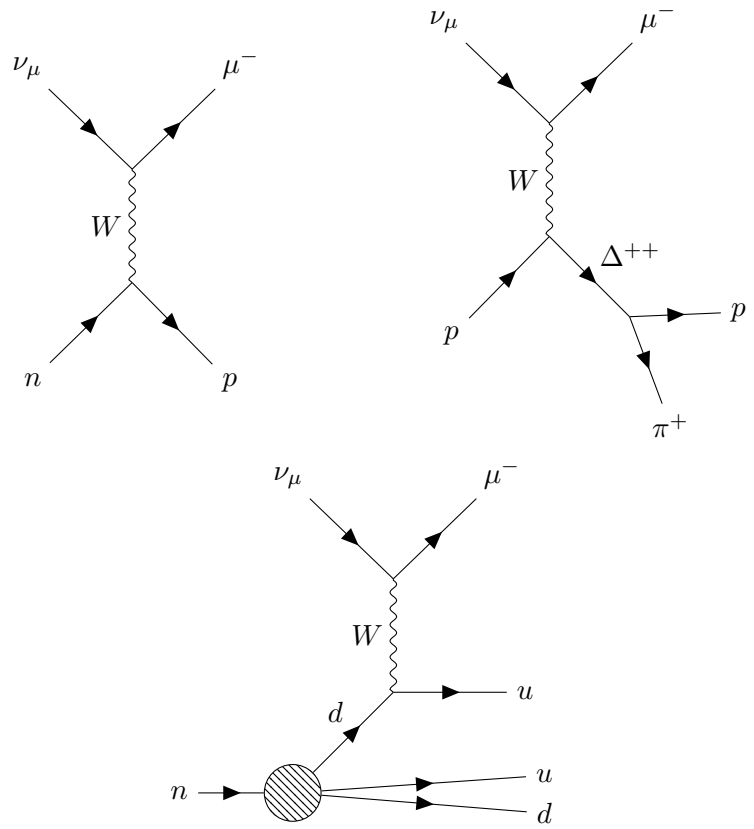


Figure 2.6: Examples of ν_μ interaction modes: quasi-elastic (top left), resonant (top right) and deep inelastic scattering (bottom). Note that the Δ^{++} baryon decays quickly to p and π^+ via the strong force, which is not shown here.

studying oscillation phenomena, the initial composition of the beam needs to be well-understood. There are several techniques for detecting neutrinos, including:

- **Scintillators:** β^+ decay can be studied using scintillation detectors. Protons in the detector medium are converted to neutrons, neutrinos and positrons; the positrons annihilate with electrons in the medium and produce detectable scintillation light. Energy thresholds tend to be low but isotropic light emission means information about particle direction is lost. KamLAND, described in Section 2.6, is one example of a scintillation neutrino detector [65].

- **Radiochemical methods:** Charged-current neutrino interactions can be used

to convert some stable atomic isotopes into unstable ones. For example, the stable isotope ^{37}Cl can be converted into the unstable ^{37}Ar, which has a half-life of about 37 days. Radioactive decays can then be used to count neutrino interactions. Experiments using this technique include the liquid gallium solar neutrino experiments SAGE and GALLEX/GNO [66, 67].

- **Bubble chambers:** Charged particles incident on an active detector volume of superheated liquid leave a trail of ionization electrons that cause the liquid to vaporize, leading to small bubbles [68]. These bubble tracks can then be optically photographed. Bubble chambers are typically subject to a magnetic field such that particle tracks curve with a radius proportional to their electric charge. One example of this was the Gargamelle bubble chamber at CERN, a freon-filled detector that operated between 1970 and 1979 and led to the discovery of neutral currents [4, 5].

- **Cherenkov detectors:** Charged particles travelling at greater than the local speed of light in a dielectric medium can produce Cherenkov light [69]. Photomultiplier tubes (PMTs) surrounding large detector volumes of suitable media (including water, ice or mineral oil) can be used to detect Cherenkov light emitted by charged leptons following neutrino interactions. Each particle creates a characteristic ring of light that can be used to infer direction and energy. Examples of neutrino Cherenkov detectors include SNO (heavy water), MiniBooNE (mineral oil), Super-Kamiokande (water) and IceCube (ice) [70–73].

- **Tracking calorimeters:** Layers of target material are interspersed with active material, allowing particles to be tracked, often with a magnetic field to measure momentum. The active material allows hadronic showers to be sampled and their energy measured. One example is the MINOS detector, comprised of steel/scintillator sampling calorimeters which collected data at Fermilab between 2005 and 2016 [74].

- **Liquid argon time-projection chambers (LArTPCs):** LArTPCs are fully-active detectors that detect ionization trails left by charged particles passing

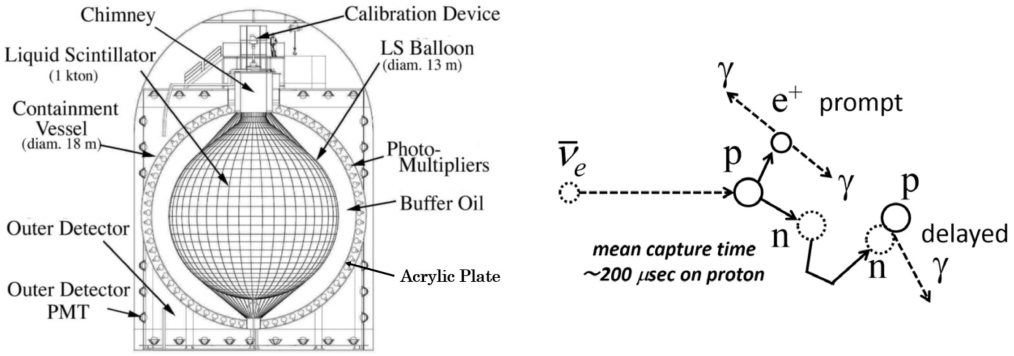


Figure 2.7: Left: a schematic of the KamLAND detector. Right: origin of prompt and delayed scintillation light from the inverse beta-decay process. Figures from [65, 78].

through the detector medium. An electric field is applied that drifts the ionization electrons towards a series of wire planes, where they are collected and measured. MicroBooNE, ICARUS and ProtoDUNE are all examples of (single-phase) LArTPCs [75–77]. An outline of the MicroBooNE detector will be given in Chapter 3.

2.6 Neutrino experiments

A number of modern neutrino experiments with various detector technologies have successfully constrained neutrino masses, measuring oscillation parameters and estimating interaction cross-sections. A selection of such experiments will be outlined here. Two future experiments, Hyper-Kamiokande and DUNE, will be described in Chapter 10 in the context of future work.

KamLAND

The Kamioka Liquid Scintillator Antineutrino Detector (KamLAND) is an electron antineutrino detector at the Kamioka Observatory near Toyama, Japan, that started collecting data in 2002. KamLAND was an oscillation experiment that studies $\bar{\nu}_e$ spectra from distant nuclear reactors to constrain oscillations parameters.

The experiment comprises a 1 kt liquid scintillator medium of paraffin oil and pseudocumene inside in a balloon-shaped container, surrounded by non-scintillating oil and

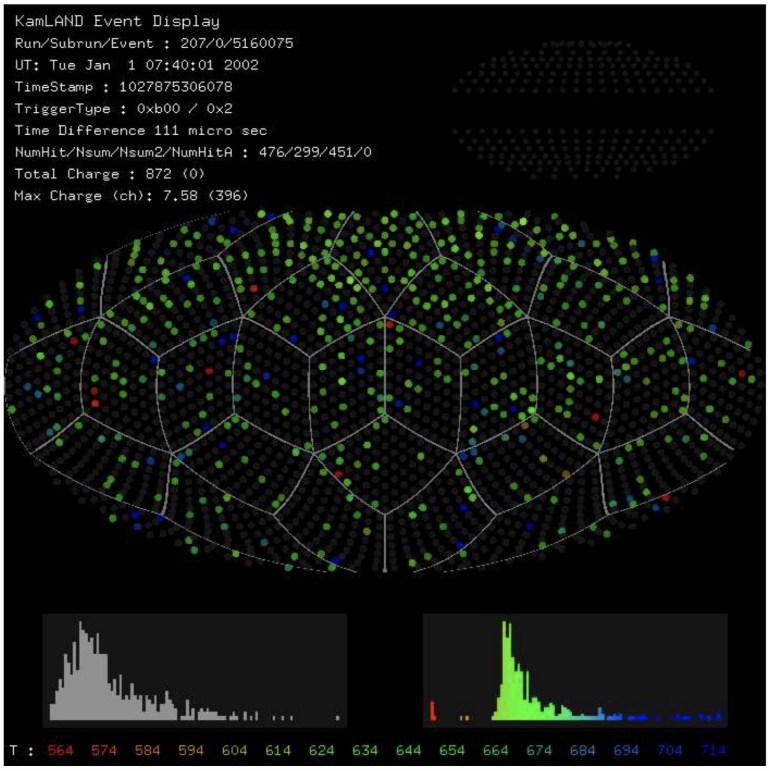


Figure 2.8: An example event display from KamLAND showing the delayed signal from a $\bar{\nu}_e$ candidate. Each dot represents the response of a PMT; the colour indicates the time at which the signal was received. Figure from [78].

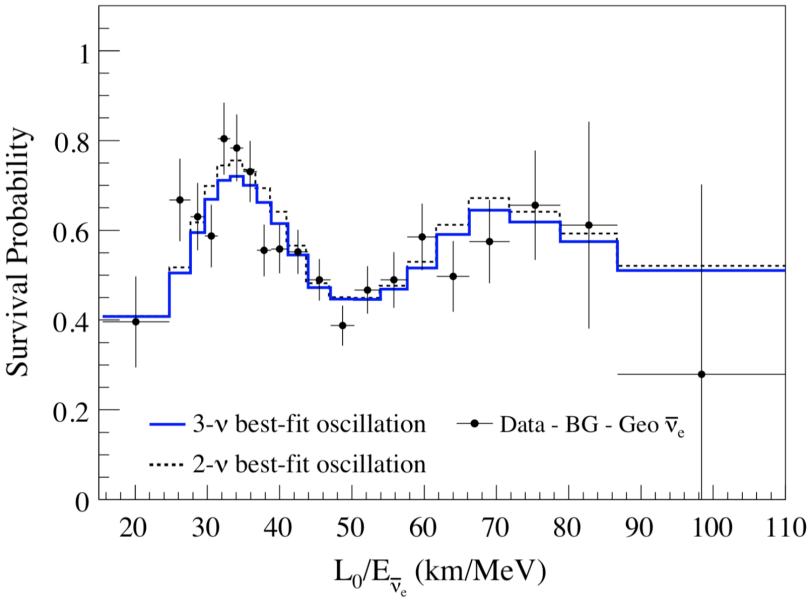


Figure 2.9: KamLAND results from 2011 showing $\bar{\nu}_e$ survival probabilities as a function of the L/E (see Equation 2.56). Survival probability here corresponds approximately to $1 - P(\nu_e \to \nu_\mu/\nu_\tau)$. This provides clear evidence for neutrino oscillations as an explanation for $\bar{\nu}_e$ disappearance. Figure from [79].

1,879 PMTs (see Figure 2.7). Incident electron antineutrinos produce prompt and delayed light through inverse beta-decay, also illustrated in Figure 2.7 (right):

$$\bar{\nu}_e + p \rightarrow e^+ + n, \tag{2.58}$$

whereafter $e^+ + e^- \rightarrow \gamma + \gamma$ produces the prompt light and neutron capture ($n + p \rightarrow d + \gamma$) produces the delayed light. PMTs detect this light (see Figure 2.8) and use it to estimate the energy of the incident neutrino: $E_\nu = E_e + E_n$. A total of 53 nuclear reactors are distributed around KamLAND, all approximately 180 km away. Since the reactors emit neutrinos isotropically, the flux from each reactor at a radius R decreases as $1/R^2$.

KamLAND has provided clear evidence of electron antineutrino oscillations that are consistent with solar neutrino experiments (see Figure 2.9) [79, 80].

Super-Kamiokande and T2K

Atmospheric neutrino results from the Super-Kamioka Neutrino Detection Experiment (Super-Kamiokande) under Mount Ikeno, Japan, provide compelling evidence of short range ν_μ disappearance, corresponding to $\nu_\mu \leftrightarrow \nu_\tau$ oscillations. Super-Kamiokande is a 50 kt water Cherenkov detector comprising a large cylindrical tank of ultrapure water surrounded by around 11,000 PMTs. An example event display is shown in Figure 2.10.

The expected ratio of ν_e to ν_μ in atmospheric neutrino spectra is fairly well-constrained. Super-Kamiokande can discriminate between electron and muon candidates by looking at the 'fuzziness' of the rings of Cherenkov light produced and therefore estimate this ratio. Differences between upward- and downward-going neutrinos can then provide evidence of oscillations over different ranges. Results from Kamiokande strongly suggest that deficits of ν_μ in atmospheric neutrino spectra are due to neutrino oscillations (see Figure 2.11) and provide strong constraints for the oscillation parameters Δm^2_{32} and $\sin^2\theta_{23}$ [82].

Tokai to Kamioka (T2K) is a long-baseline oscillation experiment that uses Super-Kamiokande as its far detector at a distance of ~ 295 km (see Figure 2.12). The near detector complex (comprising the off-axis ND280 and on-axis INGRID detectors), at a distance of ~ 280m, is a tracking calorimeter with several subdetectors for characterizing the neutrino beam. T2K has observed $\nu_\mu \leftrightarrow \nu_e$ oscillations and has more recently

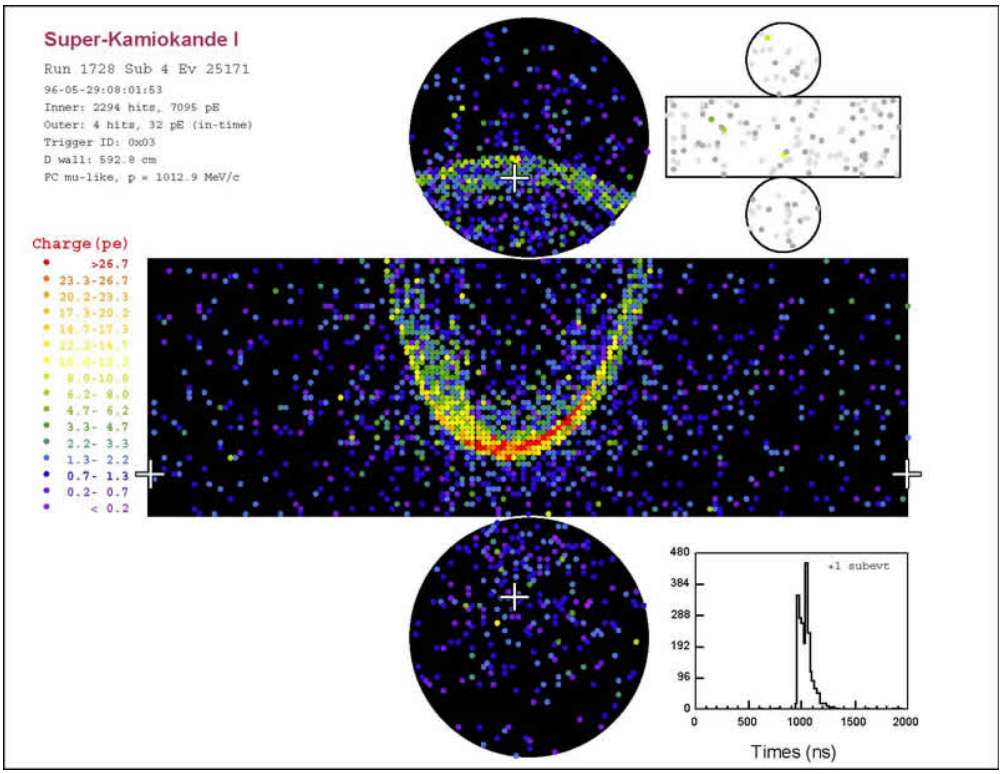


Figure 2.10: An example ν_μ event from Super-Kamiokande showing the characteristic, sharp ring of Cherenkov light produced by a muon. Figure from [81].

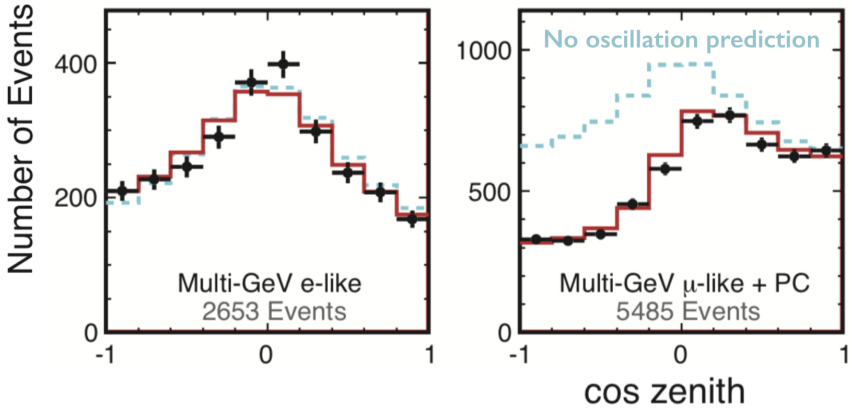


Figure 2.11: Results from Super-Kamiokande showing angular electron and muon distributions consistent with short-range $\nu_\mu \leftrightarrow \nu_\tau$ neutrino oscillations. Figure adapted from [13] and created in 2018 by the Super-Kamiokande Collaboration for the Particle Data Group.

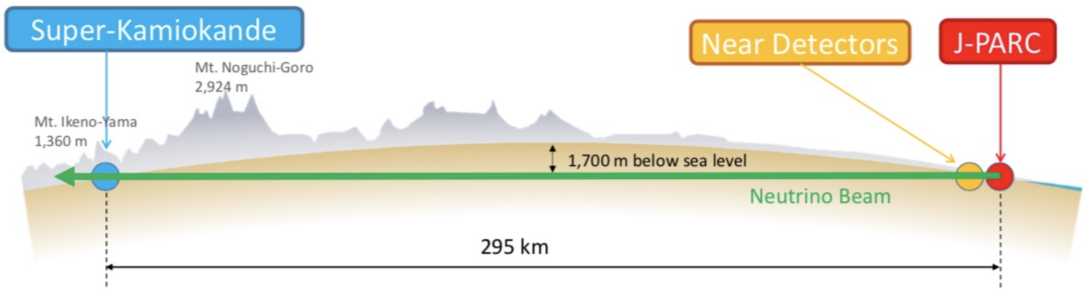


Figure 2.12: The T2K near and far (Super-Kamiokande) detectors. Figure from [83].

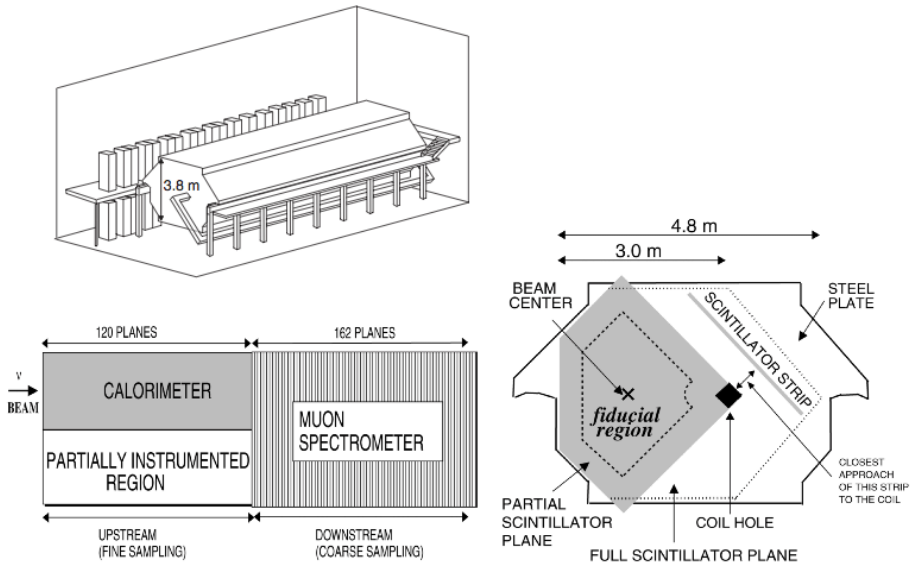


Figure 2.13: A schematic of the MINOS near detector. Figure from [86].

provided evidence at the three-sigma level that these oscillations violate CP-symmetry [84, 85].

MINOS

The Main Injector Neutrino Oscillation Search (MINOS) experiment was a long-baseline neutrino oscillation experiment in the NuMI beamline at Fermilab (described in Section 3.2). The near and far detectors were both steel/scintillator sampling calorimeters (Figure 2.13), with the 1 kt near detector at $\sim$ 1 km along the beamline and the 5.4 kt far detector at $\sim$ 735 km in the Soudan mine, Minnesota. Strips of active medium were

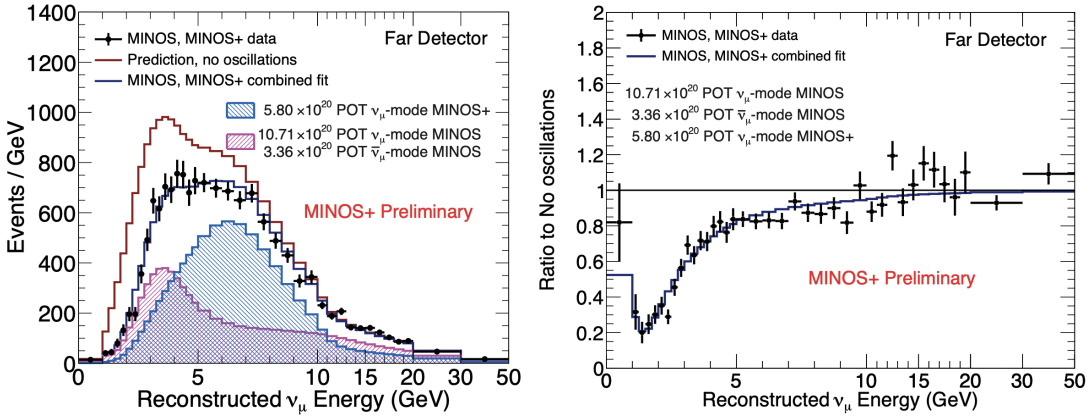


Figure 2.14: MINOS/MINOS+ oscillation results from 2017 showing clear evidence for ν_μ disappearance. Figure from [87].

placed in layers with alternating directions to allow three-dimensional reconstruction of trajectories. Topological and calorimetric signatures could then be used to identify interaction modes.

MINOS was able to further constrain the oscillation parameters Δm^2_{23} and $\sin^2\theta_{23}$ by studying the ν_μ disappearance and ν_e appearance (Figure 2.14) and perform neutrino time-of-flight calculations. In 2015, MINOS time-of-flight studies provided a value of

$$\frac{v_\nu - c}{c} = (1.0 \pm 1.1) \times 10^{-6} \tag{2.59}$$

for neutrinos in the NuMI beamline. Measuring the speed of neutrinos is an important exercise for testing special relativity and determining the mass of the neutrino. The MINOS results are consistent with both massless neutrinos and subluminal neutrinos with a small mass [88]. MINOS has also recently provided constraints on sterile neutrino mixing parameters [89, 90].

SNO

The Sudbury Neutrino Observatory (SNO) was an imaging Cherenkov detector consisting of a 1 kt container of heavy water surrounded by 9,456 PMTs (Figure 2.15) in Canada. SNO provided evidence of solar neutrino oscillations by measuring rates of

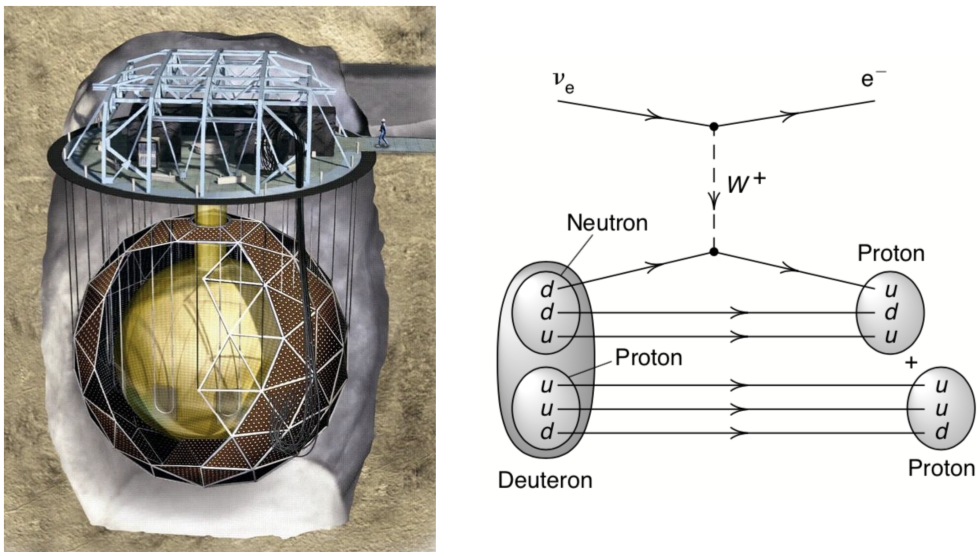


Figure 2.15: Left: a drawing of the SNO detector. Right: diagram of a typical charged-current interaction at SNO. Figures from [57].

three different interactions:

$$\begin{aligned} \nu_\alpha + e^- &\to \nu_\alpha + e^- \text{ (elastic scattering)}, \\ \nu_e + d &\to p + p + e^- \text{ (charged current)}, \\ \nu_\alpha + d &\to p + n + \nu_\alpha \text{ (neutral current)}. \end{aligned} \tag{2.60}$$

Rings consistent with different charged leptons can be used to infer the properties of incident neutrinos and cosmic ray muons (Figure 2.16). A nonzero flux of ν_μ and ν_τ was observed in the solar neutrino spectrum, along with a ν_e deficit. The flux measurements are consistent with the standard solar model (SSM), see Figure 2.17. The experiment's contribution to the discovery of neutrino oscillations was recognised by the receipt of the 2015 Nobel Prize for its director, Art McDonald. SNO is currently being upgraded to SNO+, which will search for neutrinoless double beta decay [92].

Daya Bay

The Daya Bay Reactor Neutrino Experiment is a neutrino oscillation experiment in the Guangdong province of China. The experiment targets the subdominant $\nu_e \leftrightarrow \nu_\tau$ oscillation signature using eight identically-designed liquid scintillator detectors. Its

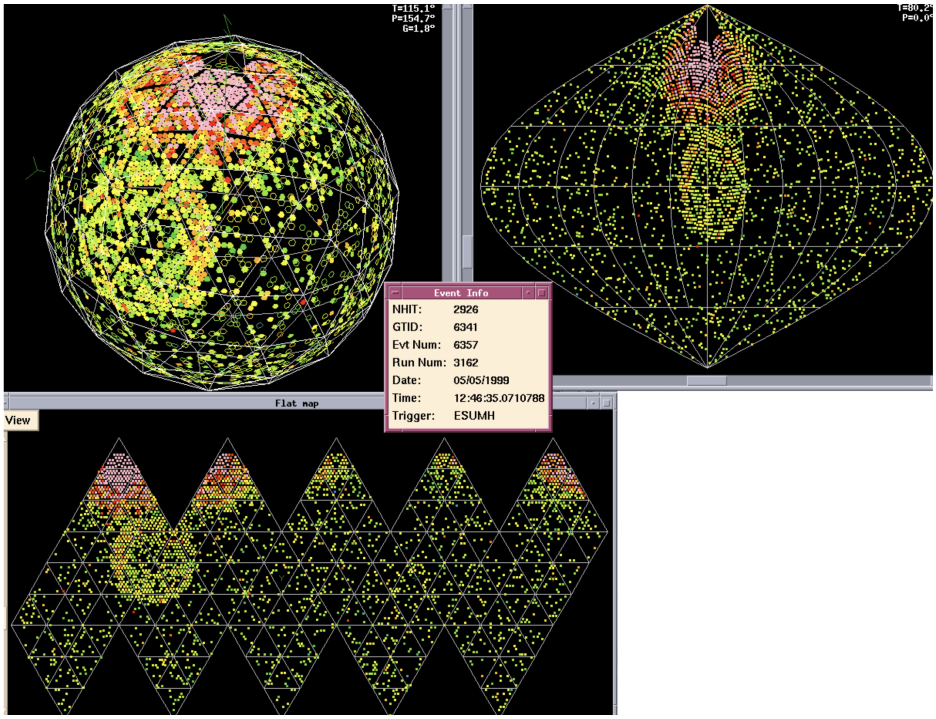


Figure 2.16: An example SNO event display. In this unusual double-ring event, a muon enterered near the top of the detector and produced the pink Cherenkov ring, then produced a secondary electron that created the lower ring. Colour represents the amount of light picked up by each PMT. Figure from [91].

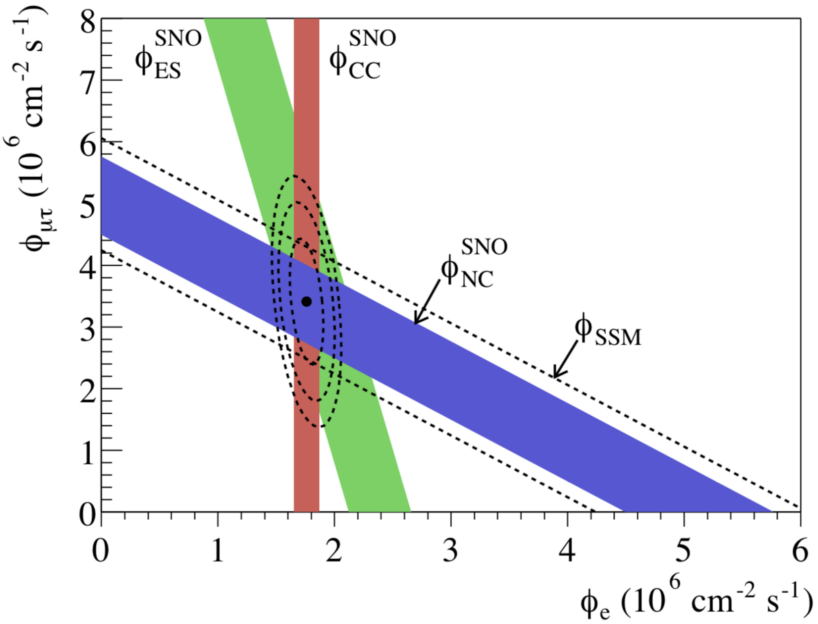


Figure 2.17: Results from the SNO experiment comparing $\phi_{\mu,\tau}$ (the flux of ν_μ and ν_τ) and ϕ_e (the flux of ν_e). The bands each represent the $\pm 1\sigma$ interval for a different interaction in Equation 2.60. Also shown is the standard solar model (SSM) prediction. Figure from [9].

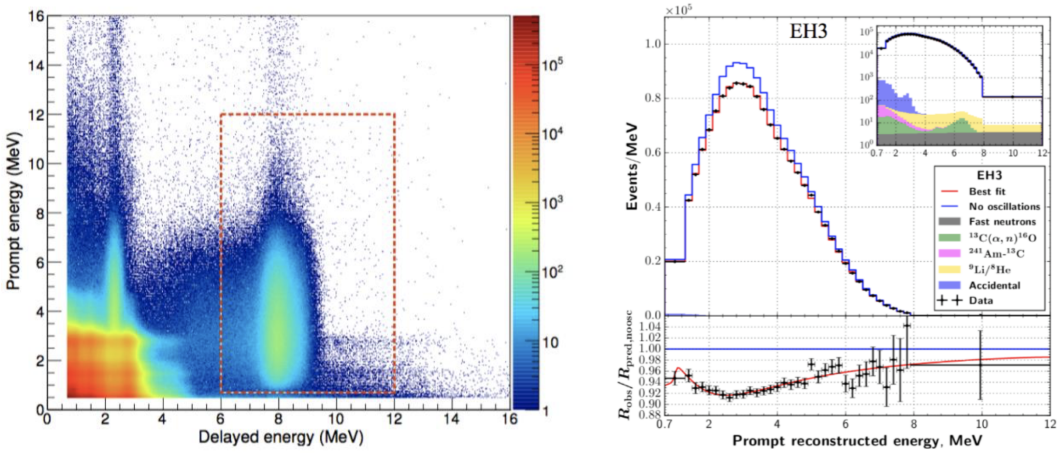


Figure 2.18: Results from the Daya Bay experiment, providing strong evidence for ν_e disappearance due to subdominant $\nu_e \leftrightarrow \nu_\tau$ oscillations. Figure from [93].

sources are the nearby Daya Bay Nuclear Power Plant and Ling Ao Nuclear Power Plant, which primarily produce electron antineutrinos.

In 2012, the Daya Bay collaboration announced a 5.2σ discovery that $\theta_{23} \neq 0$, much more significant than similar contemporary results at T2K, MINOS and Double Chooz and providing strong evidence for the subdominant oscillation signature (Figure 2.18). The experiment has since further constrained the Δm^2_{32} and $\sin^2\theta_{13}$ oscillation parameters and excluded some mass regions in the search for sterile neutrinos [94, 95], along with providing with RENO the current best measurement of the mixing angle θ_{13}.

2.7 Open questions

Despite neutrinos having been studied for many decades, there are some significant open questions in the field of neutrino physics, motivating a rich field of active research and more powerful experiments than ever before. Some of the most important questions are listed below.

- **Neutrino masses:** Whilst the squared mass differences between the neutrino mass eigenstates can be measured using existing experiments, the ordering of the masses (i.e. whether we have a 'normal hierarchy' or an 'inverted hierarchy') is still not known. T2K has provided some evidence for the normal hierarchy preference.

Oscillation experiments with significant matter effects, such as DUNE and Hyper-Kamiokande, may be able to address this question further. The absolute mass scale of the neutrinos can be constrained to some degree by existing measurements (see Table 2.1) but is still relatively unknown. Measurements from the cosmic neutrino background and beta-decay spectra may be able to further constrain the masses in the near future, for example through the recently-begun KATRIN experiment [96].

- **Oscillation parameters:** Measurement of the oscillation parameter values in Table 2.3 comes from a number of solar, atmospheric, reactor and beam neutrino experiments, including KamLAND, MINOS, SNO, T2K and Daya Bay. However, tighter constraints on these values will allow for better discrimination between theoretical models.

- **CP-violation in the leptonic sector:** To a 3σ level, no value of the CP-violating phase δ can currently be excluded (see Table 2.3). A nonzero value of δ would imply that neutrino oscillations violate CP symmetry. If $|\delta|$ is significantly larger than 0, this would be evidence in favour of the leptogenesis theory as an explanation of the matter/antimatter imbalance. Results from T2K suggest that δ is near the maximally CP-violating value $-\pi/2$ [85].

- **Cross-section measurements:** Neutrino experiment analyses rely on an accurate understanding of the neutrino interaction cross-sections. Better measurements of neutrino cross-sections in e.g. argon, carbon and water will allow experiments to reduce systematic uncertainties and place tighter bounds on other parameter measurements.

- **Majorana neutrinos:** The resolution of whether neutrinos are Dirac particles or Majorana particles is crucial for understanding the origin of neutrino mass (see Section 2.2.2). Observation of neutrinoless double-beta decay would provide strong evidence that neutrinos are Majorana particles. A number of current and future experiments aim to observe this process, including SNO+, SuperNEMO and NEXT [97–99].

- **Sterile neutrinos:** The existence of a sterile right-handed neutrino is theoretically-motivated and could help to explain nonzero neutrino mass (Section 2.2.2). Sterile neutrinos are also a candidate for dark matter. Some evidence for sterile neutrinos has been provided by LSND and MiniBooNE; MicroBooNE's main goal is to investigate these anomalies (see Section 3).

- **New physics:** Many of the current and future neutrino experiments could provide evidence for more exotic theories, such as extra dimensions and Lorentz-violating neutrino oscillations. Such new physics searches are now considered core studies in modern neutrino experiments.

Chapter 3

The MicroBooNE experiment

This chapter provides a description of the MicroBooNE experiment. It begins with an overview of its physical motivation and associated goals, followed by a summary of its neutrino sources. The MicroBooNE detector and relevant technologies are then described, including the liquid-argon time-projection chamber (LArTPC), along with a summary of signal processing techniques used to detect and reconstruct signals. The chapter ends with an overview of the techniques and software used to simulate events at MicroBooNE.

3.1 Motivation

3.1.1 The LSND and MiniBooNE anomalies

The Liquid Scintillator Neutrino Detector (LSND) was a scintillation counter experiment at the Los Alamos National Laboratory that collected data between 1993 and 1998, designed to look for evidence of neutrino oscillations. The experiment detected $\bar{\nu}_e$ interactions produced by a $\sim 800\,\text{MeV}$ $\bar{\nu}_\mu$ beam incident upon a detector medium comprising 167t of mineral oil doped with organic scintillation material. Inverse β-decay of the $\bar{\nu}_e$ produced Cherenkov and scintillation light, the timing of which (relative to subsequent neutron capture scintillation light) was used was used to detect interactions.

LSND observed a 3.8σ excess of $\bar{\nu}_e$ interactions [100]. Assuming this is the result of $\bar{\nu}_e \leftrightarrow \bar{\nu}_\mu$ oscillations would imply a large mass splitting that is incompatible with

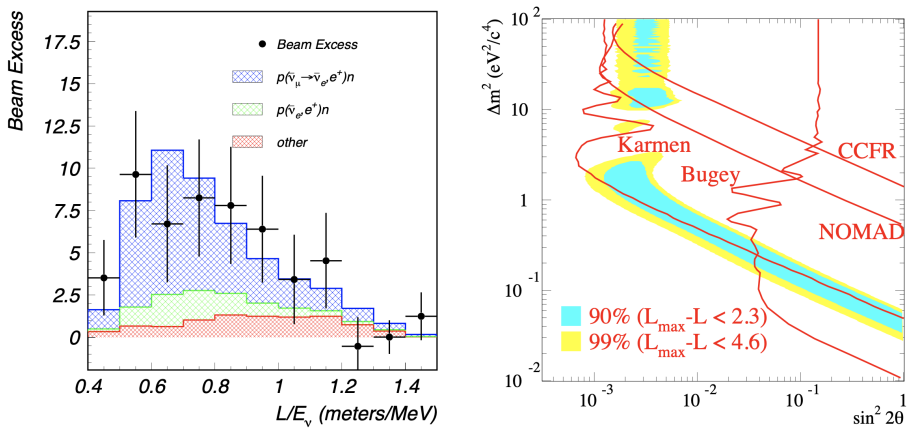


Figure 3.1: L/E_ν distribution for electron antineutrinos at LSND (left), showing the yet-unexplained excess. A theory that could explain this is the existence of a fourth neutrino, the mixing angle and mass splitting of which are constrained by the LSND observations (right). Blue and yellow regions indicate the 90% and 99% CL exclusion limits from LSND, and red lines limits from other experiments. Figures from [100].

results from other oscillation experiments. An alternative explanation that has gained traction suggests the existence of a fourth neutrino state that mixes with the three standard neutrinos in a $(3+1)$ model. Assuming that this state exists and combining LSND results with other experiments, the $(\Delta m^2, \sin^2 2\theta)$ region for these oscillations is constrained as shown in Figure 3.1.

Designed to resolve the anomaly observed at LSND, MiniBooNE is an 800t mineral oil Cherenkov detector in the BNB beamline at Fermilab (see Section 3.2) that started collecting data in 2002 [71, 101]. The experiment looks for oscillations in a similar L/E region as LSND but at a much higher beam energy ($\sim 8\,\text{GeV}$) by observing ν_e charged-current quasi-elastic (CCQE) interactions. Particle interactions in the detector medium produce Cherenkov light that is measured and characterized by the photomultiplier tubes (PMTs) that surround the medium. Disambiguation of particles is performed by the pattern of light produced: heavier particles, for instance, produce sharper Cherenkov rings, whereas neutral pion decay into a pair of photons will produce two overlapping rings. Critically, the experiment is unable to distinguish a single photon from an electron using this detection method.

The MiniBooNE Collaboration continues to analyze data. The latest result demon-

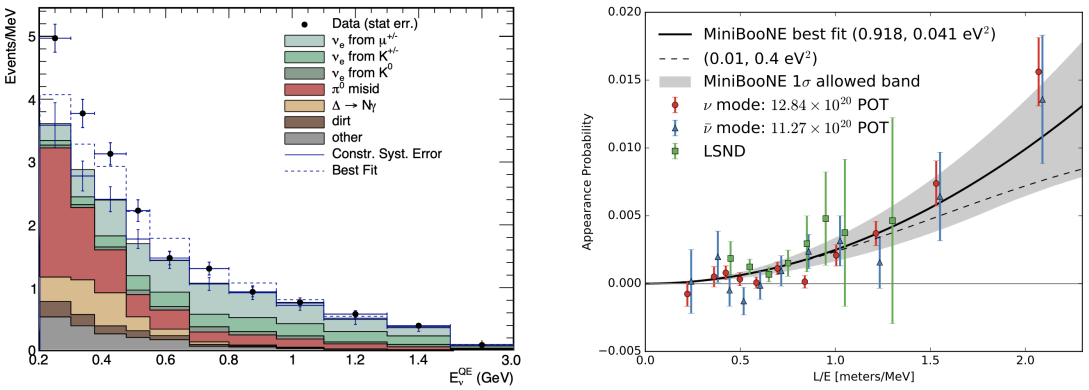


Figure 3.2: The low-energy excess of electron neutrino events observed at MiniBooNE for CCQE interactions (left). The stacked histogram represents the results expected from simulation and the points indicate the actual observations. The implied appearance probability as a function of L/E is consistent with data from LSND (right). Figures from [14].

strates a 4.7σ excess of events in a combined ν_e and $\bar{\nu}_e$ analysis [14] that is consistent with the result observed at LSND, the energy dependence of which and the combined appearance probability are shown in Figure 3.2.

There are a number of proposed explanations of the anomaly observed by LSND and MiniBooNE, including the existence of sterile neutrino states [102–104], neutrino decay [105], Lorentz/CPT violation [106] and the neutrino-photon coupling [107]. Such explanations predict either a single electron track or a single photon produced at the neutrino interaction vertex. An excess of true electrons produced from charged-current ν_e interactions could suggest BSM physics; sterile neutrino (3+1) oscillation models have found particular interest here. A detector capable of electron/photon discrimination, unlike a Cherenkov detector, is necessary to understand where the excess comes from and perhaps resolve the anomaly. Indeed, rejection of the sterile neutrino hypothesis would also be an interesting result in its own right since sterile neutrinos are one of the leading explanations of neutrino mass.

3.1.2 Physics goals

The LSND and MiniBooNE results motivated the design of MicroBooNE, the detector technology of which offers high-quality electron/photon discrimination. The physics

goals of MicroBooNE are:

- **MiniBooNE low-energy excess:** MicroBooNE's primary physics goal is to understand the source of the MiniBooNE low-energy excess. MicroBooNE therefore operates at a similar baseline distance, employs LArTPC technology (see Section 3.3) to provide electron/photon discrimination and will run for enough time to be sensitive to a MiniBooNE-type signal at a 3-5σ confidence level.

- **Cross-section measurements:** Using a liquid argon detector medium (see Section 3.3), MicroBooNE will be able to make the first measurements of exclusive final states cross-sections from neutrino scattering on argon. There is particular interest in measuring cross-sections for elastic neutrino-proton scattering, coherent pion production, and kaon and photon production in low-energy neutrino scattering.

- **Supernova detection and proton decay:** MicroBooNE primarily focuses on accelerator neutrino physics. However, it may be possible to detect supernova events in our galaxy through a short pulse of neutrino interactions measurable in the detector. Furthermore, MicroBooNE will be able to characterize the interactions of charged kaons in liquid argon, which are an important background to future proton decay searches that will take place at DUNE.

- **Testbed for DUNE technology:** DUNE makes use of similar LArTPC technology to MicroBooNE but at a much larger scale (see Chapter 10), so the much of the hardware, software and know-how developed at MicroBooNE can be translated across.

3.2 Neutrino sources

3.2.1 The Booster Neutrino Beam

The accelerator complex at Fermilab comprises four accelerators (see Figure 3.3): the Main Injector ($E_p \approx 120\,\text{GeV}$), the Linear Accelerator (Linac; $E_p \approx 400\,\text{MeV}$), the Booster ($E_p \approx 8\,\text{GeV}$) and the Recycler ($E_p \approx 8\,\text{GeV}$), together providing neutrinos

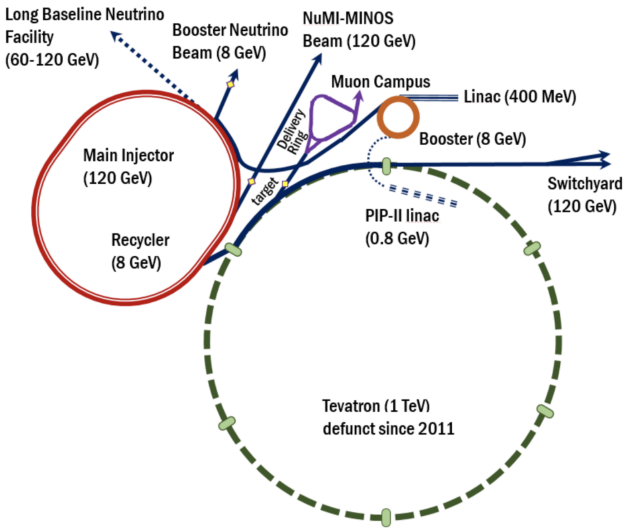


Figure 3.3: Schematic of the FNAL accelerator complex. Figure from [108].

for a range of experiments. At MicroBooNE and the wider Short Baseline Neutrino program, the main neutrino beam is the Booster Neutrino Beam (BNB) [20].

To supply the BNB, a proton beam is first created by accelerating H^- ions in the linear accelerator and stripping the electrons using a carbon foil. The Booster synchrotron accelerates these protons further and releases them as a beam spill of $\sim 4 \times 10^{12}$ protons every 1.6 μs. The beam target is a beryllium cylinder of radius 0.51 cm and length 71.1 cm (about 1.7 interaction lengths). The number of protons-on-target (POT) is measured by two upstream toroids, which measure POT to within a 2% uncertainty.

A range of hadrons are produced in the proton-beryllium interactions, which are focused using a 1.5T toroidal electromagnet called the horn. The BNB can operate in either neutrino or antineutrino mode. In the neutrino case, the horn focuses π^+ and K^+ particles, which decay into ν_μ; in the antineutrino case, π^- and K^- particles are focused. Once the desired particles have been focused, they travel through an air-filled decay pipe, then an absorber to attenuate the hadrons, leaving a highly pure beam of (anti)neutrinos, which travel through the ground before reaching the MicroBooNE detector. This is illustrated in Figure 3.4.

The BNB in neutrino/antineutrino mode primarily comprises $\nu_\mu/\bar{\nu}_\mu$ with a secondary

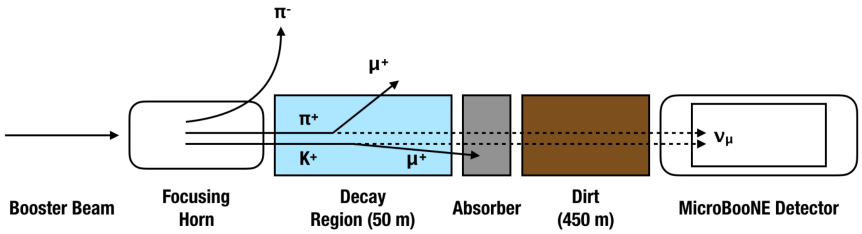


Figure 3.4: Schematic of the BNB beamline at FNAL leading up to the MicroBooNE detector. Figure from [109].

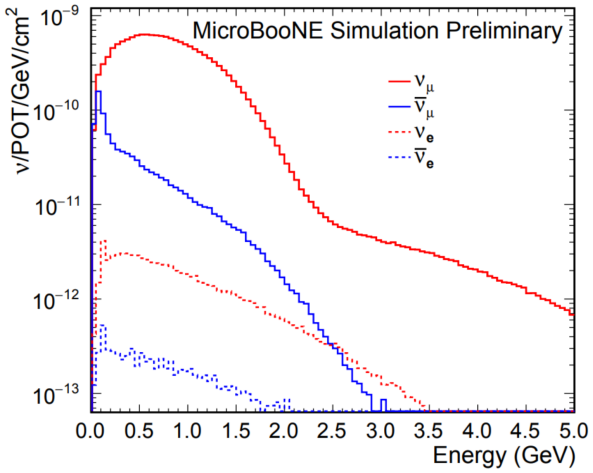


Figure 3.5: Predicted BNB neutrino flux at the position of the MicroBooNE detector in neutrino mode. Figure from [110].

contribution from $\nu_e/\bar{\nu}_e$, which result from the same muon and kaon decays. The predicted neutrino fluxes are shown in Figure 3.5. In this work, data was collected when the beam was in neutrino mode. Various decay modes lead to this flux spectrum, including the dominant $\pi^+ \to \mu^+ + \nu_\mu$ mode producing the desired ν_μ, as well as μ^+ decay ($\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$) and μ^- decay ($\mu^- \to e^- + \bar{\nu}_e + \nu_\mu$). The flux above $\sim 2.5\,\text{GeV}$ is almost exclusively due to $K^\pm$, K^0 and K^0_L decays. In total, the composition of the BNB is approximately 93.57% ν_μ, 5.86% $\bar{\nu}_\mu$, 0.52% ν_e and 0.05% $\bar{\nu}_e$.

3.2.2 The NuMI beamline

Neutrinos from the BNB are the focus of the analysis in this thesis. However, MicroBooNE also sits $\sim 8°$off-axis in the NuMI beamline and receives about $2,500$ NuMI ν_e

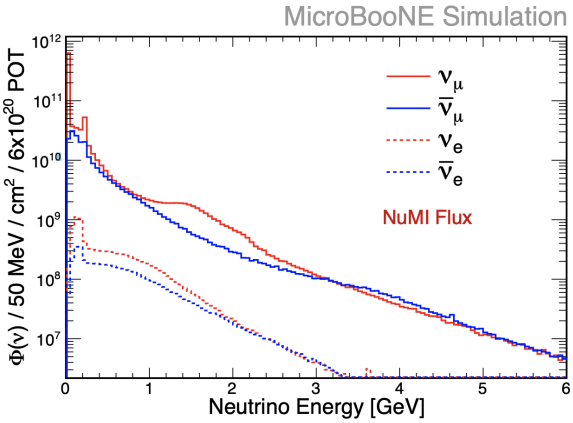


Figure 3.6: Predicted NuMI neutrino flux at the position of the MicroBooNE detector in neutrino mode. Figure from [112].

per year. The NuMI beamline operates in a similar way to the is BNB, except it is created from protons of energy $\sim 120\,\text{GeV}$ incident upon a carbon target [111]. The composition of the beam is substantially different to the BNB; its flux prediction at MicroBooNE is shown in Figure 3.6.

3.3 The MicroBooNE detector

The MicroBooNE detector comprises a rectangular cuboid LArTPC of width 256 cm, height 233 cm and length 1,037 cm set in a cylindrical cryostat and 470 m downstream of the BNB target. This is shown in Figure 3.7. This section summarises the main detector components; a fuller description can be found in [21].

3.3.1 Liquid argon time-projection chamber

A LArTPC comprises a physical volume filled with liquid argon and bounded by a cathode, an anode and a field cage. The field cage creates a uniform electric field of 273 V/cm between the cathode and the anode, which are located on two opposite long edges of the detector. The cathode consists of nine 2.3 mm-thick stainless steel sheets held at −70 kV. The anode plane, however, is a set of three offset and skewed wire planes: the induction planes U and V at ±60° to vertical and the collection plane Y at 0° to the vertical. The induction planes are each formed of 2,400 wires and the collection plane

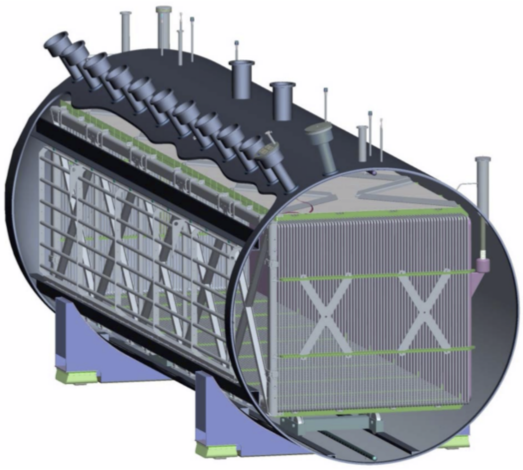


Figure 3.7: Diagram of the MicroBooNE cryostat. The wire cage can be seen as a cuboid set within the cylindrical cryostat. The Figure from [21].

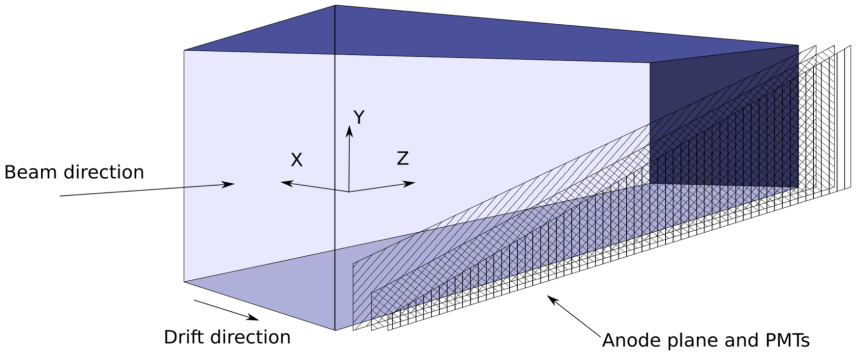


Figure 3.8: Illustration of the canonical MicroBooNE coordinate system. The cuboid indicates the boundaries of the TPC. Figure from [113].

of 3,456 wires, all separated by a pitch of 3 mm. The wires are attached to application-specific integrated circuits (ASICs) for detecting electric signals, designed to operate under cryostat conditions. Each wire plane is held at a different voltage to avoid charge being collected by the induction planes. Table 3.1 summarizes key statistics about the MicroBooNE LArTPC. The coordinate system used to describe locations within the LArTPC is shown in Figure 3.8.

Figure 3.9 illustrates how LArTPC technology detects neutrino interactions. Charged particles in the final state of the neutrino interaction leave trails of ionization electrons, which drift towards the anode plane due to the electric field. As electrons

Table 3.1: An excerpt from the MicroBooNE technical design report summarizing some key design parameters. Adapted from [21].

Key MicroBooNE design parameters	
Number of anode planes	3
Wire pitch and plane spacing	3 mm
Nominal cathode voltage	−70 kV
Nominal drift electric field	273 V/cm
Maximum drift time (nominal)	1.6 ms
Number of U wires	2,400
Number of V wires	2,400
Number of W wires	2,456
Total number of wires	8,256
U wire orientation from vertical	+60°
V wire orientation from vertical	−60°
W wire orientation from vertical	0°
U wire bias voltage	−200 V
V wire bias voltage	0 V
W wire bias voltage	+440 V

are collected on the collection (Y) plane, a unipolar signal is recorded. As they drift past the induction planes (U, V), a bipolar signal is recorded. These signals create three two-dimensional (wire number and drift time) views of an event, which are later combined to create a three-dimensional representation of the interaction.

An example of an event at MicroBooNE is shown in Figure 3.10. This display shows the view in the collection plane of a charged-current π^0 production, evidenced by two showers caused by its decay into a pair of photons. A number of muons from the cosmic ray background are similarly evident. Automatically reconstructing such an event, including taking into account information from the other two views, identifying background particles, identifying the neutrino interaction vertex, excluding noise, reconstructing particle energies and separating overlapping tracks and showers, is a significant technical challenge.

3.3.2 Light collection system

To help with both reconstruction and event selection, MicroBooNE uses a light collection system comprising 32 PMTs behind the anode plane, as shown in Figure 3.11, which is

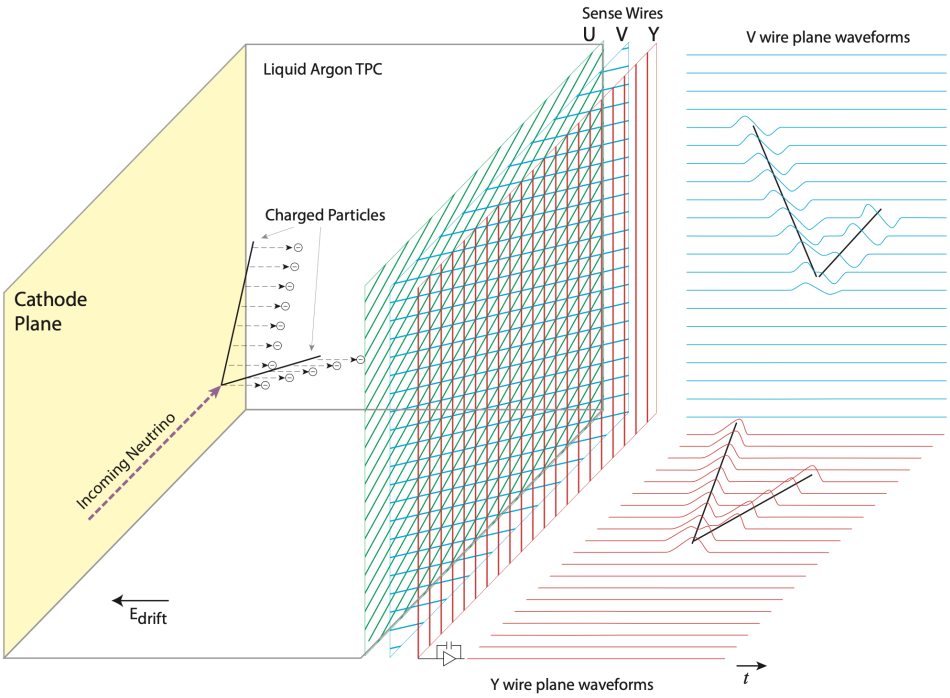


Figure 3.9: Detecting neutrino interactions with a LArTPC. The (unobserved) incident electron interacts to provide two charged particles, which leave trails of ionization electrons. These drift towards the anode plane and register signals on the collection (red) and induction (blue, green) planes. Figure from [21].

about 86% transparent to the incident light.

Light is produced in the detector via both scintillation and Cherenkov radiation. Argon is transparent to its own scintillation light, which is emitted isotropically in both prompt and delayed components during electron de-excitation, with a wavelength of about 128 nm. This wavelength is modulated to $\sim$ 425 nm using a wavelength shifter so it can be more efficiently measured by the PMTs. Coincidence of light pulses with beam spills is used for triggering, since only about 1 in 600 spills produce a neutrino interaction. This provides a significant reduction in background events.

3.3.3 Cryogenics system

The liquid argon in MicroBooNE's cryostats must be kept at a constant temperature, purity and pressure since these parameters affect the electron drift velocity and absorption length, which impacts the reconstruction of the drift time coordinate. Constancy in

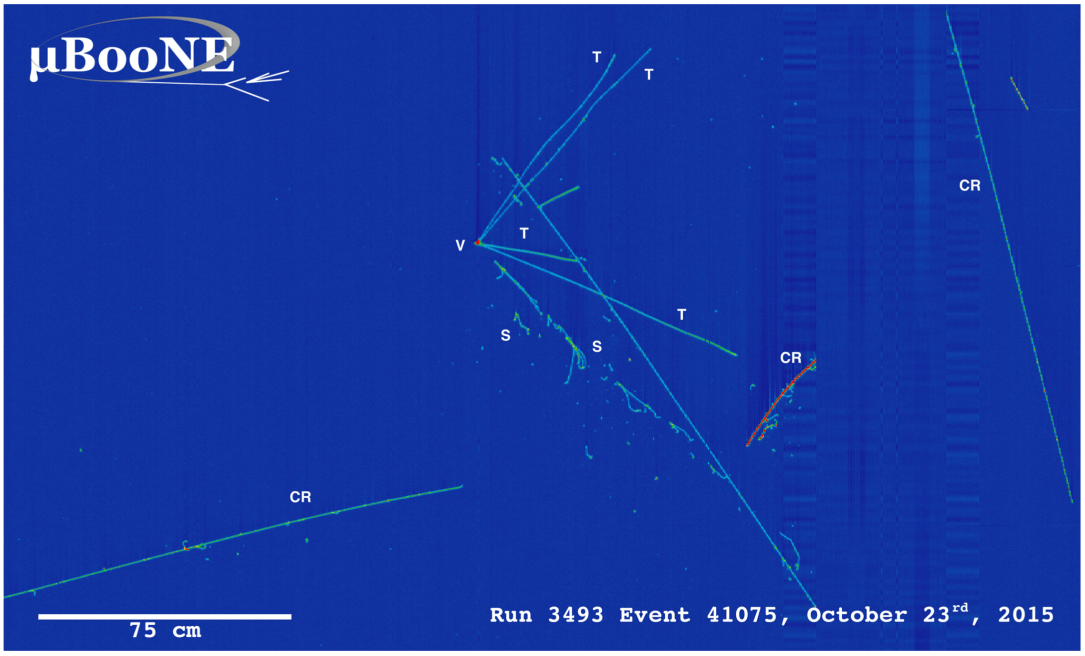


Figure 3.10: MicroBooNE event display from the Run 1 dataset, showing a candidate ν_μ charged-current (CC) π^0 interaction, along with a number of cosmic rays. The colour scale indicates the size of charge depositions at each hit. Figure from [114].

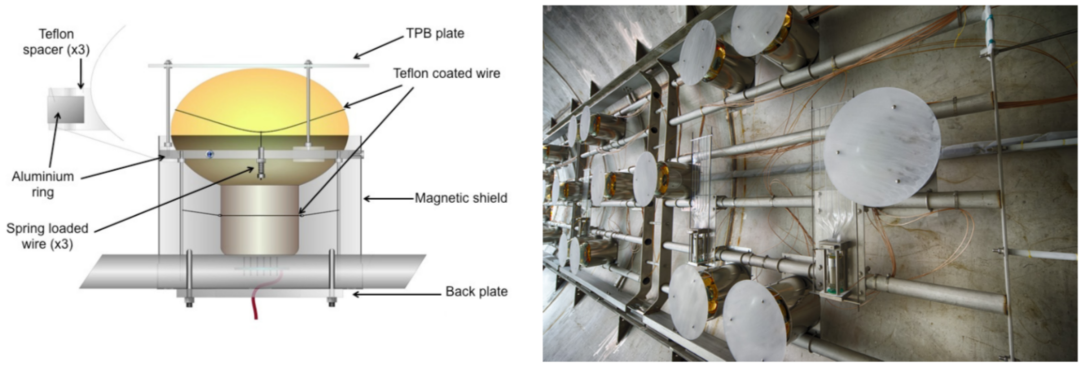


Figure 3.11: Diagram (left) and photograph (right) of the photomultiplier optical units used at MicroBooNE. Figure from [21].

time and detector position are both therefore critical. The temperature in the cryostat is monitored by 12 probes in different locations. Purification is performed using pumps and filters that circulate the detector material and remove impurities such as water and O_2 molecules.

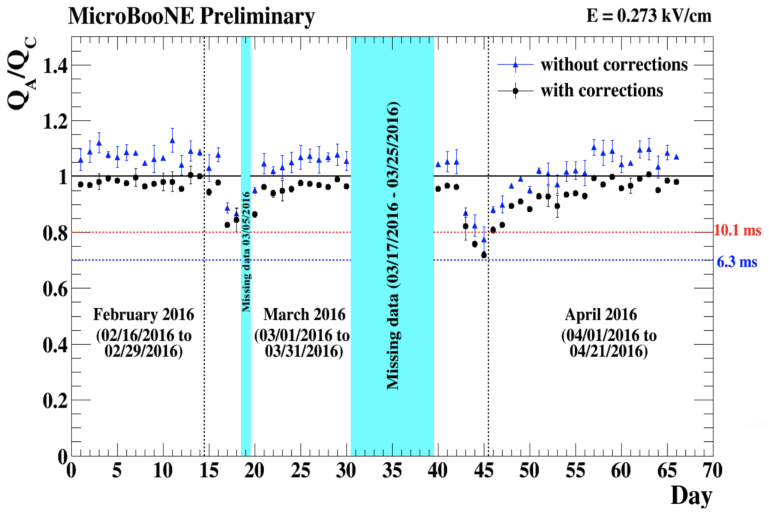


Figure 3.12: Measurements of the Q_A/Q_C ratio between February and April 2016. Variations in Q_A/Q_C stem from changes argon purity and other detector conditions. Blue bands indicate periods of missing data. The data is shown with and without space charge correction (see Section 6.4). Figure from [113].

The purity of the argon is measured via the Q_A/Q_C ratio, i.e. the fraction of electrons generated at the cathode purity monitor that reach the anode purity monitor at drift time t. This ratio is directly related to the electron lifetime τ via $Q_A/Q_C = \exp(-t/\tau)$. Multiple such monitors are placed in different positions in the detector. Variation in the Q_A/Q_C ratio is shown over a 70-day period in Figure 3.12.

3.3.4 Electronics and readout

Readout electronics at MicroBooNE can be split between those recording signals from the TPC wires and those recording signals from the PMTs, as shown in Figure 3.13. Within the TPC electronics, there exist 'cold' electronics within the cryostat that pre-amplify and shape the signal (the aforementioned ASICs) and warm electronics outside the cryostat that digitize the signal using analogue-to-digital converters (ADC) at 16 MHz before down-sampling it to 2 MHz. The outputs from each trigger are signal waveform corresponding to 9,600 ticks, equivalent to 4.8 ms. For the PMT electronics, PMT signals are amplified and digitised at 64 MHz for the software trigger and stored for later analysis. The signals are recorded over a time-window of 1,500 ticks ($\sim 23.4\,\mu$s),

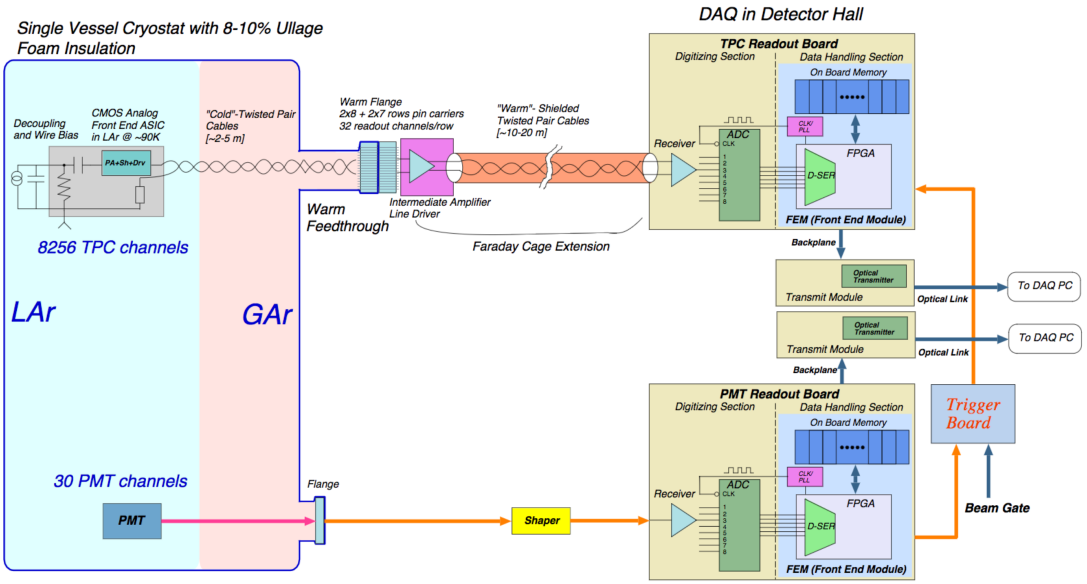


Figure 3.13: High-level schematic of the readout electronics and flow of information at MicroBooNE. Figure from [21].

starting 1.6 μs before the beam gate time.

3.3.5 Cosmic ray background and the trigger system

Being a surface detector, the rate of cosmic rays impinging upon the detector active volume is about 5.5 kHz, meaning that each drift-time window of 2.3 ms registers approximately 13 cosmic rays [115]. A cosmic-ray tagger (CRT) borders the cryostat on four sides, allowing cosmic rays to be reconstructed and tagged [116]; see Figure 3.14 for a description of the geometry and a simulation of cosmic rays. This system was installed partway through this work and therefore data making use of this tagging system was unavailable.

In order to store data parsimoniously, MicroBooNE uses both hardware and software triggers to reject windows unlikely to contain a neutrino interaction, which is the vast majority of them. Firstly, each BNB and NuMI spill triggers the hardware trigger, beginning the 4.8 ms TPC window and 23.4 μs PMT window described in described in Section 3.3.4. The BNB trigger has an efficiency of 99.8%.

Recording every such event at a 5 Hz beam spill rate would require storing $\sim$ 13 TB

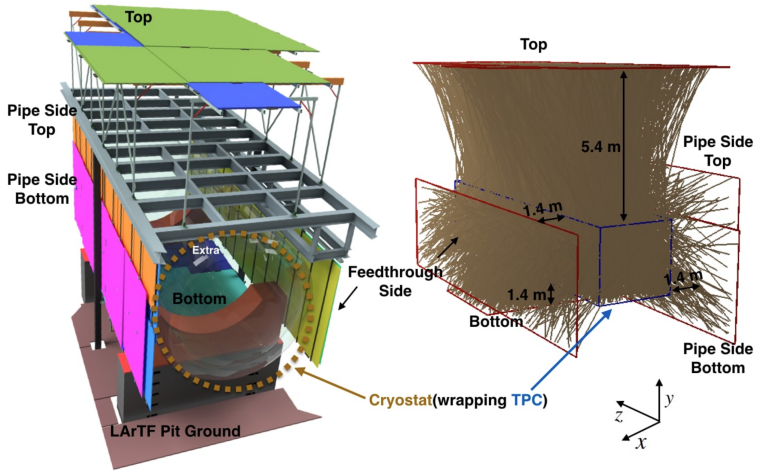


Figure 3.14: Diagram of the components (left) of the MicroBooNE muon tagging system, along with a simulation of cosmic background events interacting with the components (right). Brown line indicates simulated cosmic ray tracks. Figure from [117].

of data per day. Since so few spills produce a neutrino interaction in the detector, a software trigger is used to only record data when there is a coincident PMT trigger. This trigger allows about 3% of spills to be recorded. An additional trigger called EXT behaves similarly to the PMT trigger but only outside of the beam spill windows, ensuring that events selected by this trigger contain only background, allowing better characterization of the background.

3.4 Signal processing

Section 3.3.1 described the signals recorded by the wires in the two induction and single collection plane and induced by the drifting ionization electrons. These signals are subject to various sources of noise, which are described more fully in [118] but significant sources include $\sim 30\,\text{kHz}$ noise across all channels due to the ASICs' low-voltage regulator, noise at $\sim 36\,\text{kHz}$ and $\sim 108\,\text{kHz}$ due to the cathode high-voltage power supply, and position-dependent bursts of noise at $\sim 900\,\text{kHz}$ of unknown origin. These sources, amongst others, are deconvolved during the noise filtering process, which improves the peak signal-to-noise ratio by a factor of 3 in the induction planes and 2 in the collection planes [119]. An example of the signal on one of the induction planes

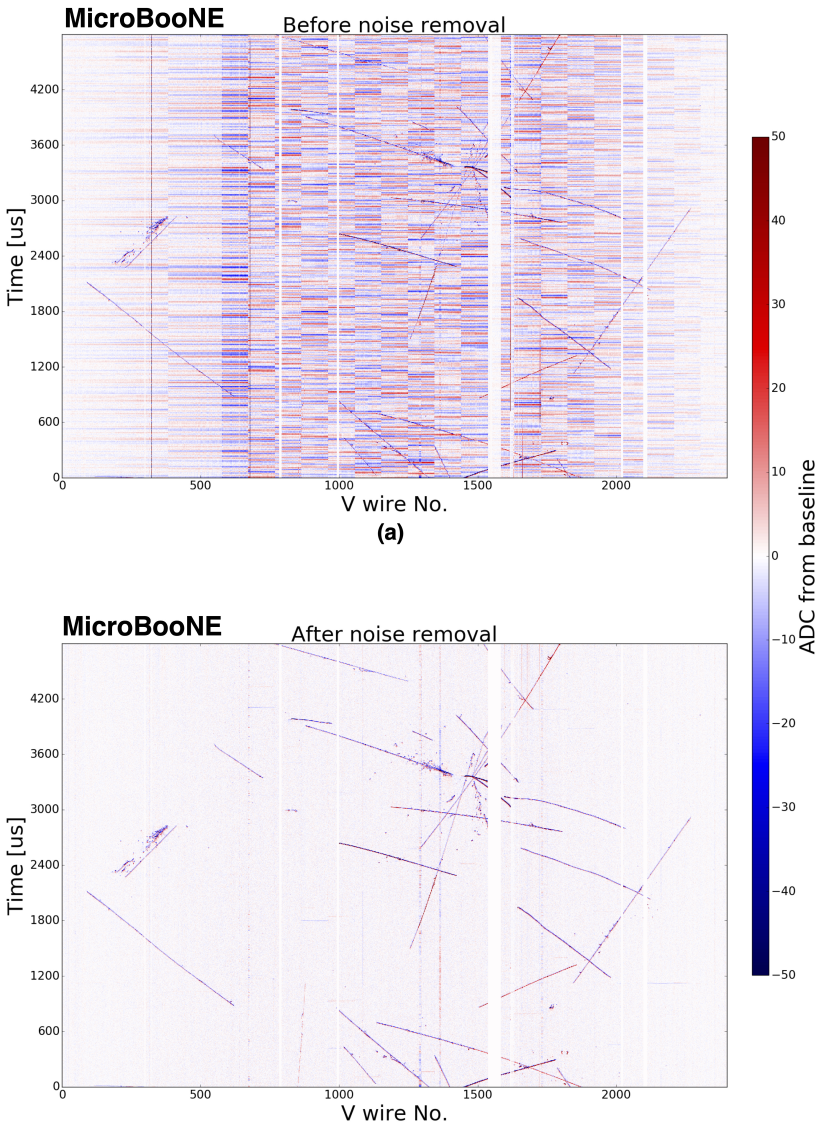


Figure 3.15: Illustration the noise removal on an induction plane (V) at MicroBooNE for a data event. The vertical axis is the drift time coordinate (x). The colour scale indicates the recorded charge depositions compared with baseline in units of ADC. Figure from [118].

before and after noise removal is given in Figure 3.15.

Finally, Gaussian fits are applied to the resulting signal to reconstruct hits, which correspond to physical regions of the detector in which charge was deposited, projected onto one of the 2D wire planes. The mean of the Gaussian corresponds to the drift time

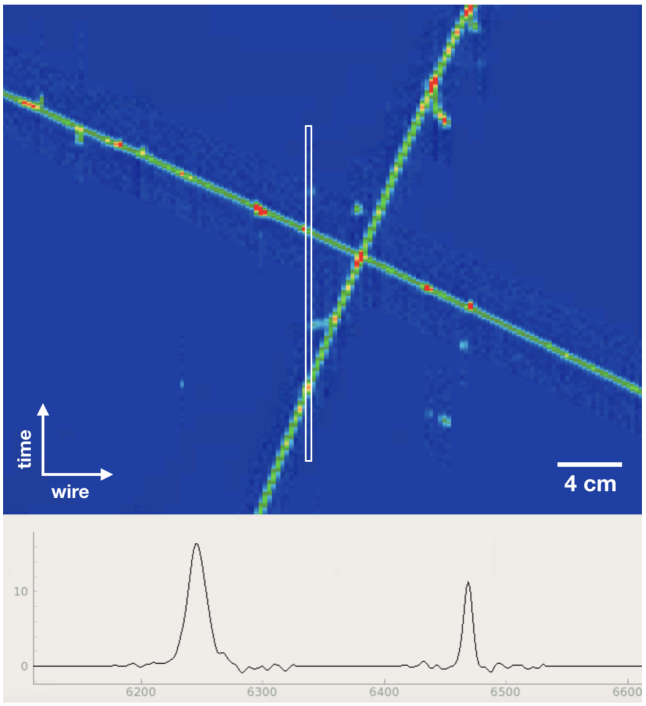


Figure 3.16: Data event display illustrating the signal recorded on one wire as a function of time. The vertical axis in the event display and the horizontal axis in the bottom figure correspond to the drift-time coordinate (x) measured in 500 ns ticks. The white box in the event display indicates the chosen wire. Figure from [112].

coordinate of the hit and its RMS to the hit width in the drift time dimension. The integral under the curve corresponds to the total charge associated with the hit. Figure 3.16 illustrates the hit reconstruction process on one wire in the collection plane.

3.5 Simulation

To facilitate physics analyses and develop reconstruction algorithms, detailed simulations of the BNB and MicroBooNE detector are under constant development. These simulations are based on Monte Carlo (MC) sampling methods. BNB flux is simulated using work done at MiniBooNE [101, 110], and events are simulated using the GENIE neutrino event generator [120] which models the neutrino interaction, including nuclear effects, branching probabilities, particle propagation through the detector and scattering of final state particles. The MicroBooNE detector is simulated using the GEANT4

software [121], which simulates particle propagation through the detector medium, the creation and motion of ionization electrons, the induction of signals on the wires, and the creation and measurement of scintillation light. The CORSIKA event generator [122] is used to simulate cosmic ray interactions.

Various GENIE, GEANT4 and CORSIKA configurations can be used to model different effects and test detector systematics; in this work, only the default MicroBooNE configurations are used. In particular, the software versions used were GENIE version 2.8.6 [120], the CORSIKA version 7.4003 [122], GEANT version 4.9.6 [123] and LArSoft version 6.26.01.10 [124].

Chapter 4

The Pandora pattern-recognition framework

This chapter introduces the Pandora pattern-recognition framework and its use at MicroBooNE. The chapter begins by outlining Pandora's motivation, philosophies and event data model, before describing the particular algorithm chains employed to reconstruct cosmic ray and neutrino interactions at MicroBooNE. The metrics used to characterize reconstruction quality and inform algorithm development are introduced, along with a summary of the Pandora reconstruction quality at MicroBooNE.

4.1 Introduction

Following the signal processing stage described in Section 3.4, three sets of 2D hits with associated positions, extents in the wire-number and drift-time coordinates, and associated integrated ADC counts (relating to the amount of charge deposited in the detector at that position) are available for a given event. To perform physics analyses, these products must first be transformed into physically meaningful 3D tracks and showers; a process known as reconstruction. Solving this pattern-recognition problem is the remit of Pandora [125].

The Pandora software is divided into a software development kit (SDK) and experiment-specific libraries. The SDK defines the experiment-agnostic functionality

for performing reconstruction, including the event data model (EDM), application programming interfaces (APIs) of general use across experiments, and a C++ framework for object lifetime management and for developing chains of algorithms that conform to the Pandora algorithm paradigms. The experiment-specific libraries define specific chains of algorithms that reconstruct events at a given experiment, including the definition of experiment-specific objects.

Pandora is developed under the multi-algorithm philosophy; in other words, that a reconstruction chain should consist of a large number of relatively simple algorithms, each of which addresses a specific event topology or performs a parsimonious task to assist with the reconstruction. Layers of such algorithms slowly build up a full picture of the event, rather than relying on one complex, monolithic algorithm that attempts to cover all topologies at once. Decisions made within each algorithm are conservative to avoid making mistakes early in the reconstruction that are difficult to undo. This allows for a highly complex reconstruction process to be reduced to a set of digestible, maintainable and reusable algorithms.

4.2 Event data model

Pandora's event data model (EDM) is a hierarchical framework of objects with parent-daughter relationships and associations, designed in such a way that each object contains all the information required to fully specify its relationship and interaction with all other objects in the EDM without repetition. Pandora considers input objects to be immutable, whereas algorithm objects are created within and iterated on through algorithms. Algorithm objects may be associated with the immutable input objects, but the Pandora SDK's object lifetime management framework ensures that such a relationship is unique, such that input objects are not double-counted by algorithm objects. The Pandora EDM is illustrated in Figure 4.1. The Pandora input objects are:

- **CaloHit:** The 2D hits, each from one of the three wire planes, that form the primary input to Pandora. Each one has a physical position, spatial extent (in the wire-number and shared drift-time dimensions) and an associated charge-like measurement, recorded as the integrated ADC count.

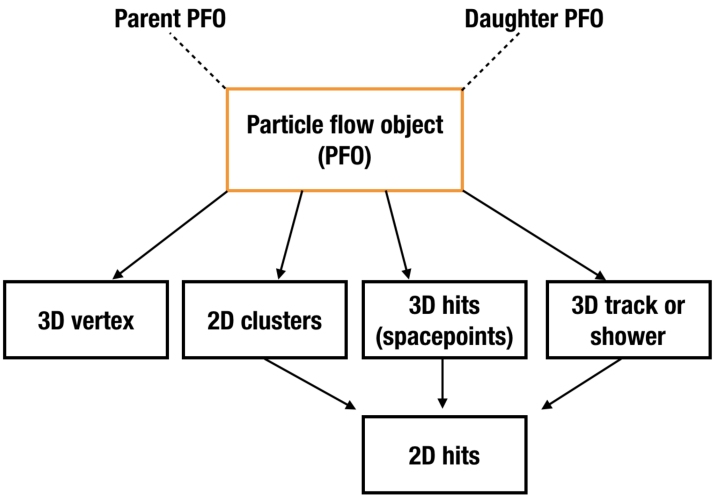


Figure 4.1: The Pandora event data model (EDM). These are the outputs of the pattern-recognition process. Solid lines indicate associations between objects; dash lines indicate the navigable particle hierarchy (see main text). Figure from [109].

- **Monte Carlo particle:** A set of information produced by the particle generator that, following detector simulation, led to the observed hits. This includes the particle type, hierarchy and kinematic properties. These objects are used to help debug and develop reconstruction algorithms and are only available for simulated events.

Some of the Pandora algorithm objects are:

- **Cluster:** A collection of CaloHits, typically used as intermediate products in the process of creating tracks or showers. There is a 2D version of this object, which can contain CaloHits only from one view, and a 3D version, which can contain 2D clusters from multiple views.

- **Vertex:** A spatial position in the detector indicating a place where an interaction or decay occurred. This object can be 2D, in which case it can be considered the projection of a spatial position onto a wire plane, or 3D, in which case it fully specifies a point in space.

- **Track/shower:** A 3D object containing one or more clusters that should correspond to a whole particle or electromagnetic shower.

- **Particle flow object:** A hierarchical object that can contain clusters, vertices and tracks/showers, allowing navigation through the parent-daughter hierarchy that encodes the temporal order of interactions. These form the complete final output of Pandora, fully and parsimoniously defining the relationship between objects in the Pandora EDM.

Modular algorithms are run in a predetermined order which can create, modify, associate and destroy the algorithm objects described above. These algorithm chains are not hard-coded, but instead defined externally using XML settings files, which also define the algorithm configuration if required. More detail about the Pandora SDK and EDM can be found in [125].

4.3 Reconstruction at MicroBooNE

Pandora is the default software used for pattern recognition at MicroBooNE, where the fine spatial resolution requires the reconstruction of complex, often overlapping topologies in the presence of significant noise. Pandora integrates with Fermilab's liquid argon reconstruction software LArSoft [124] via a module that translates between the Pandora EDM and the LArSoft EDM. The previously described input objects are passed and translated from LArSoft into Pandora, along with detector geometry (the dimensions of the active volume, information about wire plane angles and pitches, and dead detector regions). The output PFO hierarchy is translated into LArSoft tracks and showers and additional processing such as Kalman filtering is performed inside LArSoft.

Reconstruction at MicroBooNE is twofold: the first reconstruction chain is `PandoraCosmic`, which is optimized to reconstruct cosmic ray muons. One such optimization is that the vertex of each particle can reliably be taken to be the hit with the highest y coordinate. Following this chain, a cosmic ray tagging module is run on the reconstructed output and tags particles which are highly likely to be cosmic rays, based on the track trajectory and beam trigger timing. The hits corresponding to these cosmic rays are removed before running a second algorithm chain, `PandoraNu`, on the remaining hits, which comprise ambiguous cosmic rays and potentially a neutrino interaction. This chain is optimized for reconstructing neutrino interactions, such as first

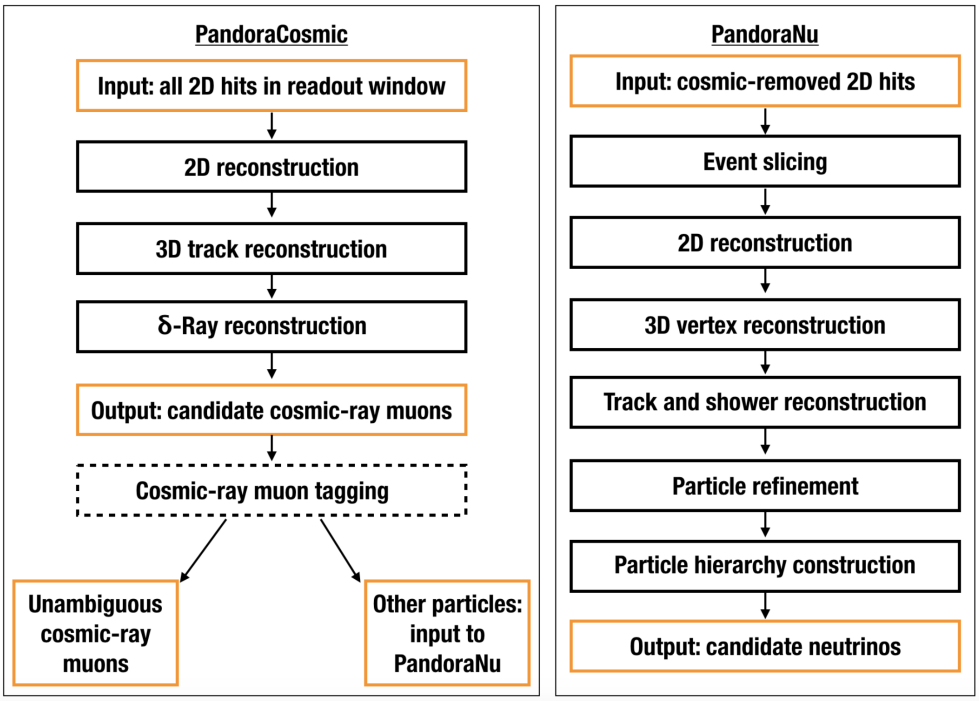


Figure 4.2: The Pandora algorithm chains for cosmic rays reconstruction and neutrino reconstruction, as performed at MicroBooNE. Orange-bordered boxes indicate inputs and outputs of Pandora. Figure from [109].

'slicing' the event into cosmic ray and neutrino interaction candidates and employing more sophisticated algorithms for selecting the interaction vertex. Every slice is reconstructed under both cosmic and neutrino hypotheses, i.e. using each algorithm chain. The twofold reconstruction chain is illustrated in Figure 4.2.

4.4 Pandora algorithm chain

This section will describe some of the reconstruction algorithms that are used at MicroBooNE and elsewhere. We will exclude the crucial neutrino interaction vertex reconstruction algorithms, which are described in detail in Chapter 5.

4.4.1 Two-dimensional clustering

Figure 4.3 illustrates part of the 2D clustering process. First, clusters are created by grouping hits together based on their physical proximity and linearity of the fit passing

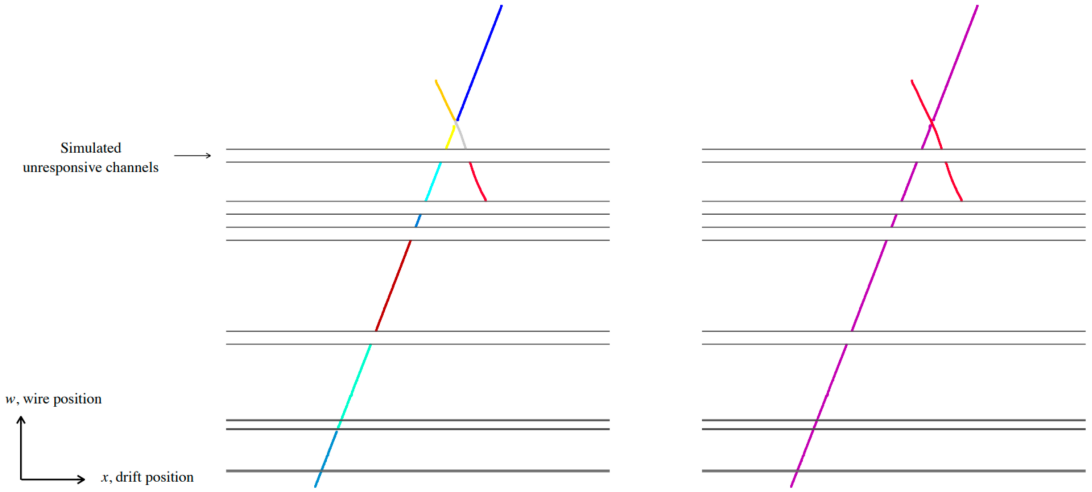


Figure 4.3: Illustration of the 2D hit clustering process for two crossing tracks in the presence of simulated unresponsive channels, indicated by empty detector regions bordered by grey lines. Different colours indicate different clusters before the gap information is used to merge the clusters (left) and afterwards (right). Figure from [126].

through them. Any ambiguity here prompts the end of a cluster, resulting in a large number of small clusters. This process is error-averse, resulting in high-purity but low-completeness clusters, in the sense that the hits corresponding to each true particle are typically split into many distinct clusters but ones with low contamination from other particles.

Following this, clusters are merged by considering their spatial proximity and the alignment of their principal axes, extrapolating cluster trajectories through detector gaps. These algorithms operate only on 2D hits, such that it is independently performed three times for each event (once for each wire plane). The output is a set of 2D clusters, each containing a set of hits ideally the complete set of induced by a different particle for each cluster.

4.4.2 Track reconstruction

The purpose of the track reconstruction algorithm set is to reconcile the 2D clusters in each view that correspond to track-like particles by splitting, merging and matching them to produce 3D tracks. There are various distinct topological cases with different signatures in each view; in concordance with the Pandora multi-algorithm philosophy,

each class of topology is addressed by a different track reconstruction algorithm that targets only that class without disrupting the others.

A metric quantifying the goodness of match between each cluster in each of the three views is stored in a 3D tensor, representing every possible cluster combination. Clusters, which form the indices of the tensor, can be merged or split by algorithms, resulting in an updated tensor. The tensor with the minimum matching ambiguity can then be selected as the final set of matches. To construct the matching metric, the shared x (drift-time) coordinate in each view is exploited by projecting a set of common x-sampling points from every pair of views to each third view following a sliding least-squares local linear fit. For a given set of matched clusters, repeating this process for each pair and comparing the projected positions with the observed positions yields a χ^2-like parameter which is used to quantify the match quality, along with the x-overlap span and the number of consistent sampling points.

Some of the track reconstruction algorithms[1] that address different topological patterns are listed below. During track reconstruction, these are run in the order listed below; if any of the algorithms makes a change to the tensor, then the entire chain is run again. This process is iterated until the whole chain is run with no changes being made to the tensor. Illustrations of some particular cases are given in Figure 4.4.

- **Clear tracks:** An unambiguous topology in which there is exactly one cluster captured in the given x-extend in each view. This is shown in the top left in Figure 4.4.

- **Long tracks:** A topology in which small clusters that present matching ambiguities are rejected in favour of longer tracks. This is shown in the top right in Figure 4.4.

- **Overshoot tracks:** A topology in which two views present two clusters and one view presents one cluster, typically when the projection in one view taken alone makes the pair of clusters look like a single cluster. The resolution here is to split the cluster in the single-cluster view. This is shown in the bottom left in Figure 4.4.

[1]These are in fact *algorithm tools* under the Pandora SDK definition, since they are reusable algorithm-like objects that can be called from Pandora algorithms.

- **Undershoot tracks:** A topology in which two views present one cluster and one view presents two clusters, for instance when missing hits lead to a cluster being erroneously split. This is shown in the bottom right in Figure 4.4.
- **Missing tracks:** If a cluster is entirely missing from one view, which is possible for example due to detector gaps, missing hits or overlapping particles in a given projection, tracks can be made from clusters in only two views.
- **Track splitting and missing track segment:** These algorithms address topologies in which hits at the end or in the middle of clusters are missing, or where clusters need to be split for better matching.
- **Long tracks (reduced threshold):** This algorithm is run again with an easier-to-achieve matching threshold in an attempt to match clusters that remain outstanding.

Clusters that end this process unmatched are dissolved and re-clustered; frequently these correspond to fragments of δ-ray showers. Following this, 3D hits are reconstructed in a topology-dependent way through techniques to select the y and z coordinates that minimize χ^2.

4.4.3 Shower reconstruction

The `PandoraNu` pass looks to reconstruct electromagnetic showers which may be produced downstream of the neutrino interaction due to electrons and photons. An example of such a simulated event is given in Figure 4.5. This process begins in 2D, where shower-like clusters are identified using a support vector machine (SVM) trained on MC simulation to classify track versus shower clusters. Long 2D clusters identified are shower-like, especially those that point back to the neutrino interaction vertex, are considered to be shower spines. Full 2D showers are then recursively 'grown' from the spines by incorporating 'branch' clusters based on their proximity and relation to the existing shower.

Following 2D shower reconstruction, 3D showers are created by matching the 2D showers between views using a tensor method similar to the track reconstruction algorithm. In this case, edges of the 2D envelopes fitted to the showers are projected rather

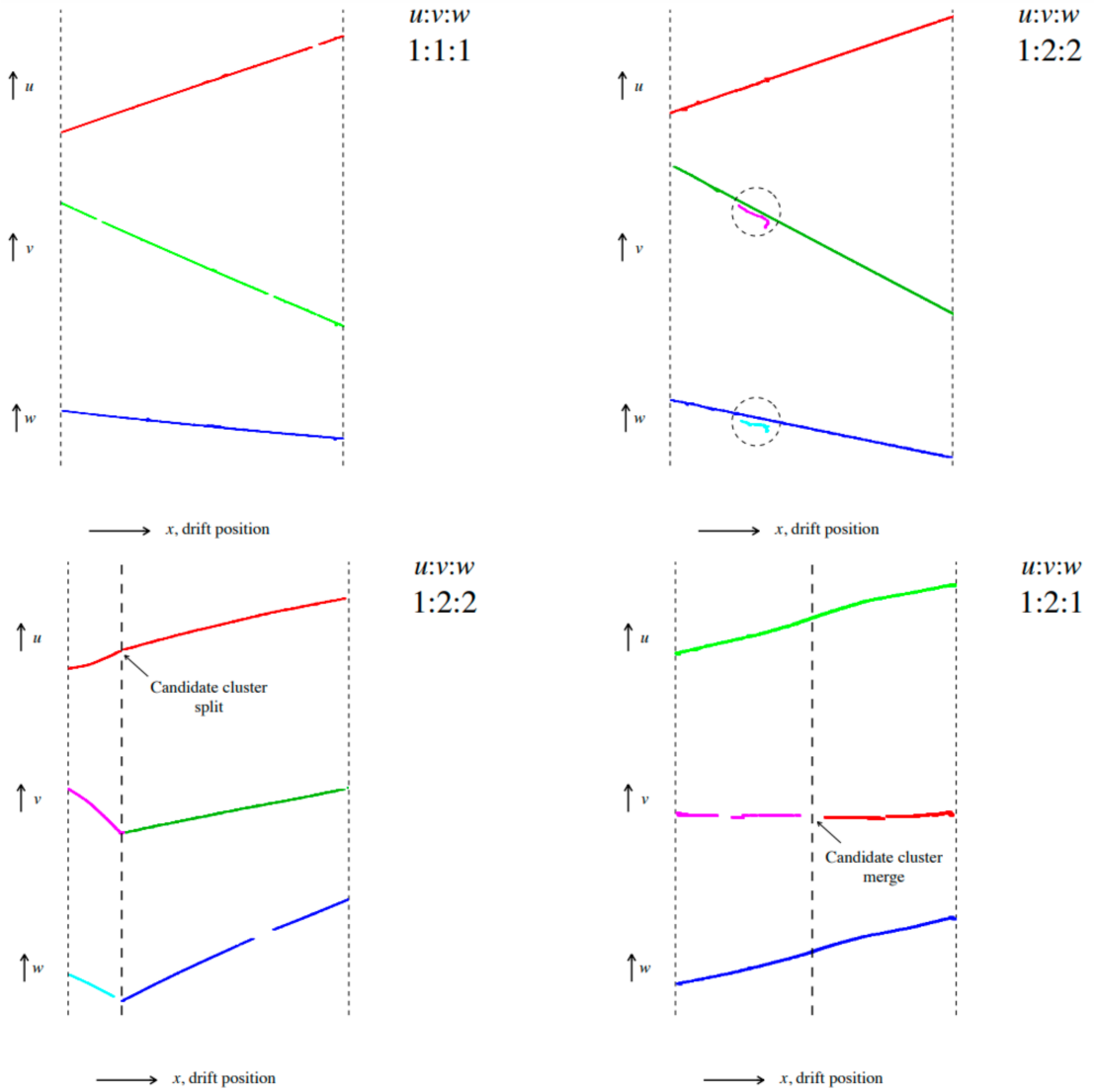


Figure 4.4: Illustration of the 'matching' process between 2D planes to create 3D tracks in four different cases. Note that the drift-time coordinate (x) is common to all planes. Top left: an unambiguous single cluster in each plane. Top right: the presence of a δ ray is inferred from the V and W planes despite it being missing in the U plane. Bottom left: information from the V and W planes is used to split an erroneously merged cluster pair in the U plane. Bottom right: information from the U and W planes is used to merge and erroneously split cluster in the V plane. Figure from [126].

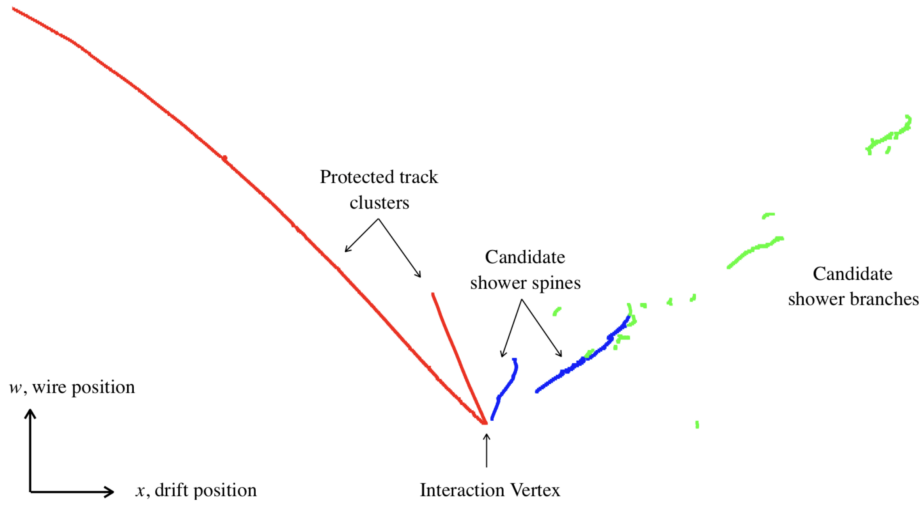


Figure 4.5: The electromagnetic shower reconstruction process in the W plane. The red track-like clusters are protected from alteration at this stage. The blue clusters are tagged as candidate shower spines and the green clusters as candidate shower branches. The branches are iteratively added to the spines to 'grow' high-quality showers. Figure from [126].

than sliding linear fits to the hits themselves. The metric of goodness of fit is then the proportion of hits in the third view that are contained within the projected envelope, as well as the x-overlap details.

Finally, an SVM-based algorithm is used to classify every particle as either track-like or shower-like based on topological and calorimetric properties; this flag is not used during reconstruction but is made available downstream for analysis.

4.4.4 Particle refinement and hierarchy building

The final broad step in the reconstruction chain is refinement of reconstructed particles and hierarchy building. This includes further growing of shower-like particles using unassociated clusters through extrapolating cone fits in both 2D and 3D, and by considering the proximity of unassigned hits to existing showers. The results is a set of track-like and shower-like particles for each detector slice. This is used to create the particle hierarchy: a set of parent-daughter relations between particles with the (unobserved) neutrino particle as the root node. Based on the proximity of a particle's 3D hits to the current interaction vertex, the hierarchy is grown from the neutrino inter-

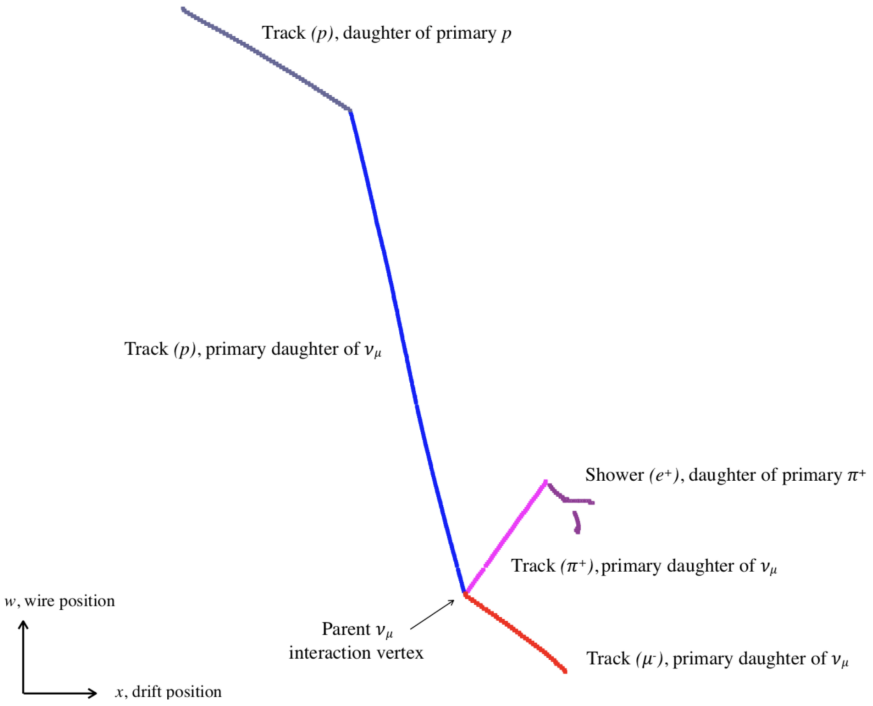


Figure 4.6: Example particle hierarchy for a CC ν_μ event with a muon (red), proton (blue) and charged pion (magenta) in its final state. The final-state particles are considered to be daughters of the (unseen) neutrino, and objects downstream of each final-state particle are considered daughters of that particle; e.g. the shower (dark purple) at the end of the charged pion track becomes a daughter of the charged pion. Figure from [126].

action vertex outwards. Unassigned particles are assigned as daughters of the primary daughters of the neutrino, using a number of algorithms that each address a different topology. An example of such a hierarchy is shown in Figure 4.6.

4.5 Performance

Quantification of reconstruction quality is crucial for algorithm development. In Pandora, a set of reconstruction metrics have been developed based on the matching of hits into 'true' particles using MC information. Each hit is tagged by the particle which contributed most of its charge; for each MC particle i, this forms the set of 'true hits' T_i. Similarly, the actual hits associated with some reconstructed particle j form a set

Table 4.1: Pandora reconstruction performance for three selected channels, as measured by the number of reconstructed particles matched to each true (MC simulated) particle. The correct value of 1 reconstructed particle per true particle is indicated in bold. Adapted from [126].

		Number of matched particles			
Channel	Particle	0 (%)	1 (%)	2 (%)	3+ (%)
CCQE $\nu_\mu + \mathrm{Ar} \to \mu + p$	μ	1.3	**95.8**	2.9	0.1
	p	8.9	**87.3**	3.6	0.2
CCRES $\nu_\mu + \mathrm{Ar} \to \mu + \pi^+ + p$	μ	3.5	**95.1**	1.4	0.0
	π^+	9.0	**86.8**	4.0	0.3
	p	6.9	**80.9**	11.4	0.8
CCRES $\nu_\mu + \mathrm{Ar} \to \mu + \pi^0 + p$	μ	3.7	**94.8**	1.5	0.0
	p	9.9	**85.5**	4.3	0.3
	γ_1	6.8	**88.0**	4.8	0.4
	γ_2	29.9	**66.4**	3.6	0.2

R_j. A detailed account of the performance metrics and performance at MicroBooNE can be found in [126]; this will be summarized here.

Only particles deemed 'reconstructable' by a specific definition are included in performance metrics so that particles that are too small to be reliably reconstructed, as well as sparse cloud of hits caused by neutrons, do not unfairly influence the metrics. The definition excludes hits downstream of a far-travelling primary photon, downstream of a far-travelling neutron, particles producing fewer than 15 high-quality hits (where high-quality here means that at least 90% of the energy recorded by the hit was deposited by the particle under consideration), and those which have fewer than six high-quality hits in at least two views.

To be able to construct meaningful metrics, reconstructable true particles then need to be matched to reconstructed particles. This process is performed by first uniquely matching each MC particle to the reconstructed particle with which it shares the most hits; once every MC particle that shares hits is uniquely matched, the remaining reconstructed particles are assigned to the MC particle with which they share the most hits. This means that each MC particle may be matched with no, one or more than one reconstructed particle. The metrics used to evaluate reconstruction quality are listed below:

- **Purity** $P_i := |T_i \cap R_i| \,/\, |R_i|$**:** For a given match, the fraction of shared hits that come from the true particle.

- **Completeness** $C_i := |T_i \cap R_i| \,/\, |T_i|$**:** For a given match, the fraction of true hits that are present in the shared hits.

- **Efficiency** η**:** For a given true particle type, the fraction of such particles with at least one reconstructed particles matched to it.

- **Correct event fraction:** The fraction of events for which each true (reconstructable) primary particle has exactly one reconstructed particle matched to it. The results displayed for three channels in Table 4.1 illustrate this process.

Per-particle reconstruction metrics are typically reported only for each primary particle in the final state of the neutrino interaction (i.e. each primary daughter of the neutrino) since these are most important to physics analyses and Pandora folds daughters hierarchically back to primary particle, though it is possible to construct these metrics for any particles in the event. Reconstruction efficiency for three channels as a function of kinematic and reconstruction parameters is shown in Figure 4.7.

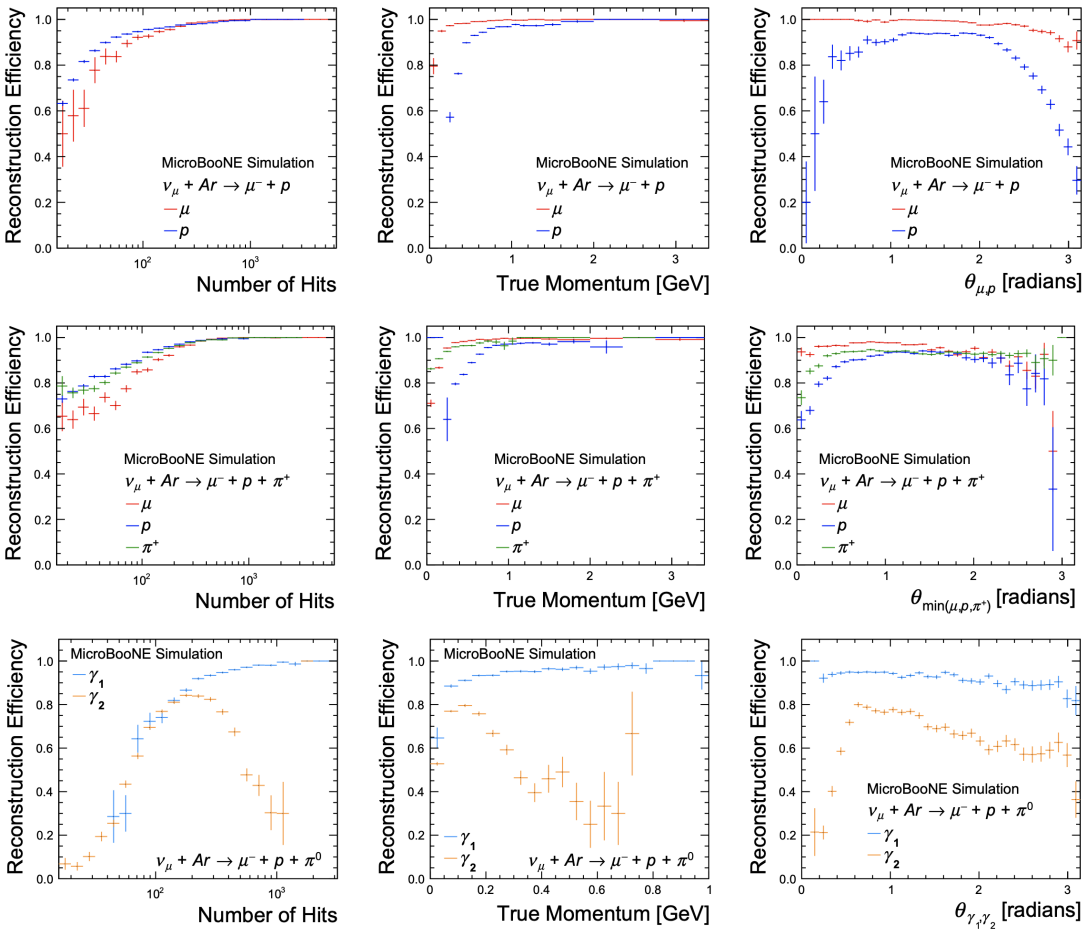


Figure 4.7: Per-particle Pandora reconstruction efficiencies as a function of the number of reconstructed 3D hits for each particle (left column), the true momentum of each particle (middle column) and the opening angle between the particles (right column) for three channels. The top row shows the $\nu_\mu + \mathrm{Ar} \rightarrow \mu^- + p$ channel, the middle row the $\nu_\mu + \mathrm{Ar} \rightarrow \mu^- + p + \pi^+$ channel and the bottom row the $\nu_\mu + \mathrm{Ar} \rightarrow \mu^- + p + \pi^0$ channel. Figure from [126].

Chapter 5

Neutrino interaction vertex reconstruction

This chapter presents a new algorithm for neutrino interaction vertex reconstruction, which is now the default algorithm used at MicroBooNE and has enabled a number of new physics results. The features and models that are used in the new algorithm are described, along with the previous vertex algorithm used in Pandora. The performance of the algorithm and its effect on reconstruction quality is evaluated on per-beam and -channel bases.

5.1 Introduction

Overall neutrino event reconstruction quality is tightly coupled to the quality of the vertex reconstruction. As described in Chapter 4, vertex reconstruction occurs early in the 2D reconstruction and its result is used to split/merge 2D clusters and inform the structure of the resulting particle hierarchy. The placement of the vertex has far-reaching consequences both to downstream reconstruction and to physics analyses making use of these vertices. Misplaced vertices can lead to inappropriate track/shower splitting or merging, which in turn lead to misconfigured particle hierarchies, backwards-going particles, poor particle identification (due to altered calorimetric profiles) and poor energy reconstruction. Establishing the neutrino interaction vertex is therefore one of

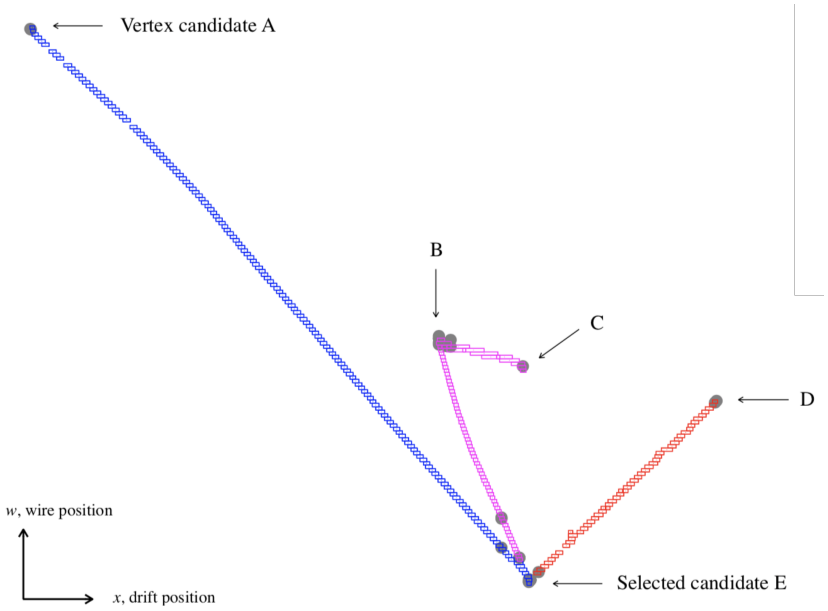


Figure 5.1: Illustration of a typical configuration of candidates in a simple simulated event. The figure shows 2D clusters and projected 3D vertex candidates in the W view. The w direction is equivalent to the z (roughly beamline) direction at MicroBooNE, and x is the drift time coordinate common to all 2D views. Figure from [126].

the most important tasks performed during reconstruction.

In Pandora, vertex reconstruction is split into two distinct tasks:

- **Vertex candidate generation:** The creation of a large number of 3D vertex candidates based on features of the event, typically a few hundred. In practice, this is achieved by identifying interesting event features in one or more 2D planes, such as ends of tracks/showers and kinks in tracks, and matching them in other views. The goal is to ensure that a good candidate is very likely to exist amongst the many. Vertex candidates must correspond to reconstructed hit positions in at least one 2D plane. Note that the vertex candidate may lie in an unresponsive region of the detector in some planes. This is illustrated in Figure 5.1.

- **Vertex selection:** The ranking of all vertex candidates based on their plausibility and selection of the most likely one. The information available to make the selection at this point in the reconstruction includes: the list of 3D vertex candidates (which can be projected into any 2D view); preliminary 2D clusters, which

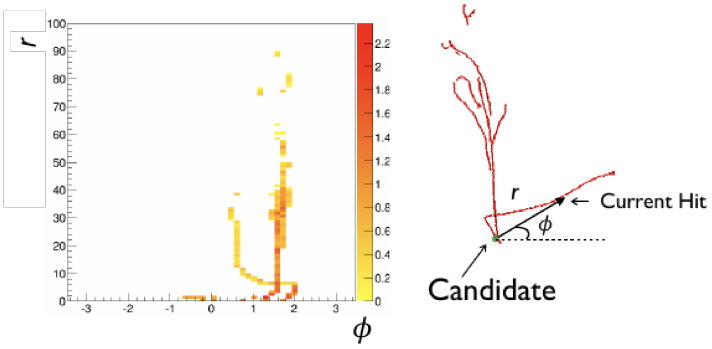


Figure 5.2: Illustration of the `RPhi` algorithm. The position of each nearby hit is represented in 2D polar coordinates and the peakiness of this distribution with respect to the ϕ axis is characterised to make a score, with more distant (i.e. higher r) hits' contributions deweighted.

are labelled as track-like or shower-like;[1] charge depositions associated with each 2D hit in each cluster; and, 2D positions of each hit in relation to the detector geometry.

This chapter concerns a new algorithm developed for the vertex selection process, improving upon the previous algorithm by adding new features inspired both by physical intuition and issues noted at MicroBooNE. The resulting feature set is then used as input to a machine learning (ML) model, the first of a number of ML models now employed inside Pandora. The algorithm described here has enabled a number of physics results at MicroBooNE, including a track multiplicity study [127] that examined events with up to 6 particles in the neutrino interaction final state.

5.1.1 The `RPhi` algorithm

The previous algorithm used for vertex selection in Pandora is the `RPhi` algorithm, which assesses vertex quality using the implied geometry of the hit positions local to each candidate. The score considers 2D polar coordinates centered on the candidate and characterizes the straightness of the pattern of emerging hits. Algorithmically, the 2D histogram shown in Figure 5.2 is projected onto the ϕ axis with each hit's contribution

[1]Note that this cluster labelling is used internally in Pandora to assist with the reconstruction. A version tailored for downstream analysis that takes into account the information available at the end of the reconstruction has been developed, based on the work presented in this chapter.

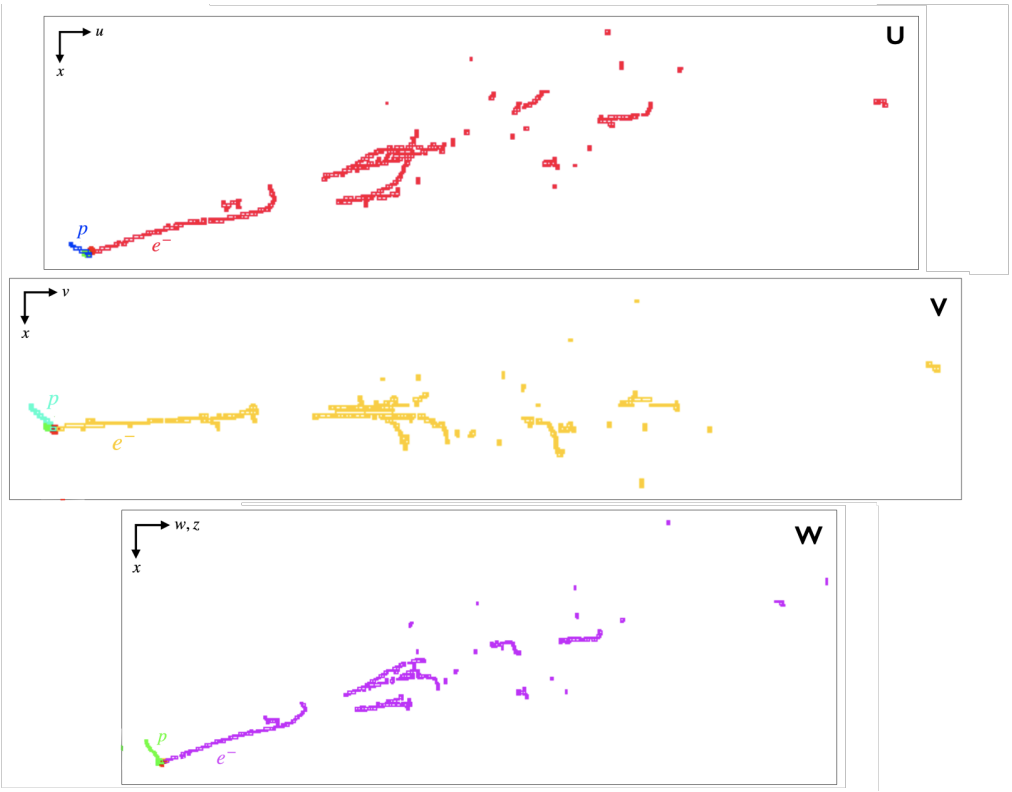


Figure 5.3: Illustrative simulated intrinsic ν_e ($e^- + p$) event. The true vertex (which need not lie on a hit) is indicated in each view by a green circle and a reconstructed vertex by a red circle.

weighted by $1/r$ to create a 1D histogram. This histogram is then turned into a single scalar score by characterizing the peakiness of the distribution, where a more peaky distribution indicates a better candidate. There are three ways in which the peakiness can be calculated; in the simplest version, the histogram is normalized and the score is the sum of the squares of the values in each bin. Physically, more plausible candidates are more likely to be 'pointed to' by downstream clusters, especially by clusters in the vertex's immediate vicinity as their particles are less likely to have scattered or undergone secondary interactions. The resulting score is then modified by the 'beam deweighting' factor f_{deweight}, which takes into account the fact that candidates at lower values of z, the coordinate roughly aligned with the beam direction, are more probable due to final state products being more likely to travel closer to the beam direction than

backwards:

$$f_{\text{deweight}} := e^{-z'/\zeta}, \tag{5.1}$$

where

- $z' := (z - z_{\text{min}})/(z_{\text{max}} - z_{\text{min}})$ is the modified beamline coordinate;
- z is the candidate coordinate along the longitudinal detector axis (roughly aligned with the beam direction);
- z_{min} and z_{max} are the minimum and maximum longitudinal coordinates of the vertex candidates; and,
- ζ is a tunable constant.

Figure 5.1 shows an example where maximizing the beam deweighting factor alone would lead to the optimal vertex choice; Figure 5.3 shows an example where it would not. This process is repeated in each 2D plane and the candidates receive an averaged overall score, which is used to rank the candidates and select the best one. Note that the algorithm takes into account only geometric information, neglecting calorimetric information. This will be used as the baseline algorithm for assessing the performance of the new vertex algorithm developed in this chapter.

5.2 Feature building

We improve on the algorithm presented in the last section by incorporating a variety of new pieces of information into a model, including calorimetric data, event-level features and additional geometric context. We take the beam deweighting score and the score from the `RPhi` algorithm, which we now call the 'r-ϕ score'. We additionally add a number of new features, which are described below.

Energy kick score

In order to incorporate calorimetric information into the vertex selection process, we introduced the energy kick score. The energy kick score is akin to the summed transverse

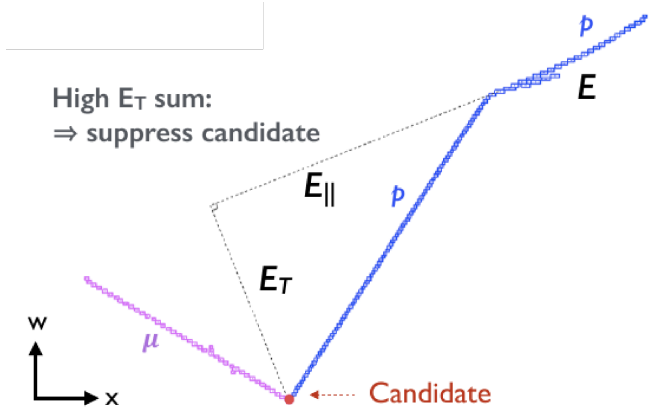


Figure 5.4: Illustration of the energy kick score, which resembles the 'transverse charge' of each cluster with respect to the candidate such that candidates with a high energy kick score should be suppressed. The basis of this intuition is that candidates which seem to require a larger amount/number of visible energy 'kicks' to explain the observed direction of downstream particles are less likely to be the correct candidate.

energies of the clusters with respect to the vertex candidate. Physically, particles in the neutrino interaction final state are likely to point back towards the true neutrino vertex and, whilst particles in the final state of downstream interactions are less likely to do so, they also tend to be less energetic. In practice, we use charge rather than energy to avoid performing calibration; at this stage, 2D clusters are not particularly well defined. The score is defined as

$$S_{\text{kick}} := \sum_{\text{view } i} \sum_{\text{cluster } j} \frac{Q_j(x_{ij} + \delta_x)}{d_{ij} + \delta_d}, \tag{5.2}$$

where

- Q_j is the summed charge of the hits of cluster j;
- x_{ij} is the transverse impact parameter between the candidate and a linear sliding fit to cluster j in view i, extrapolating linearly from the end of the cluster if need be;
- d_{ij} is the closest approach between the candidate and cluster j in view i; and
- δ_x and δ_d are empirical, tunable constants that control the behaviour of the score, particularly for very close clusters; we set them as 10 cm and 0.06 cm, respectively.

We note that setting δ_x and δ_d to 0 corresponds to a true 'transverse charge' measurement but behaviour when $d_{ij} = 0$ is then undefined, which is not an uncommon case since each candidate must lie on a hit in at least one view. This calculation is illustrated in Figure 5.4.

Local asymmetry score

The local asymmetry score measures whether the energy deposited locally to the candidate falls more isotropically or more on one side. A high degree of local asymmetry makes it less likely that the candidate has been placed in the middle of a track, a common case when the track is broken in one view or has undergone a secondary interaction (e.g. muon δ-rays or proton scattering). We only consider the case of one or two local clusters and otherwise issue the maximum asymmetry score of 1. This is because there is no local asymmetry information in a view where there are no local clusters and, when there are 3 or more local clusters, the axis about which the asymmetry is defined can become less meaningful and lead to spurious results. We also note that observing three or more clusters close to a vertex often indicates a favourable candidate anyway. The steps to calculate the local asymmetry score are outlined below:

1. Consider the candidate in the U projection.

2. Find all the track-like 2D clusters with at least one hit within 5 cm of the candidate. If there are 0 or more than 2 clusters, return the maximum allowed score of 1 (i.e. no suppression). Perform a sliding linear fit to these local clusters.

3. Calculate the 'local event axis' by performing a charge-weighted sum of the cluster directions at the hit closest to the candidate. If there are two clusters, consider the local cluster direction opening angle θ_{12}; if $\cos\theta_{12} < -0.9962$, then rotate one of the directions by 180° before performing the sum. This is to ensure a meaningful local event axis in the common case of candidates lying in the middle of straight tracks.

4. Project each hit from the local clusters onto the local event axis.

5. Sum the charge of the hits whose projections fall on the left side E_{L} and the right side E_{R} of the candidate. Calculate the local asymmetry score in this view: $S_{\mathrm{LA}} := |E_{\mathrm{L}} - E_{\mathrm{R}}|/(E_{\mathrm{L}} + E_{\mathrm{R}})$.

6. Repeat this procedure in the V and W projections, and return the sum the results.

This score provides useful information but it cannot provide a high-quality vertex selection in isolation due to invisible products, the 2D projection and calorimetric reconstruction quality. In Figure 5.1, for instance, candidates A, D and C will receive asymmetry scores near 1 as they are close to the ends of clusters. Candidate E will receive a score of 1 as it lies at the intersection of more than two tracks. The candidate partway down the magenta track is likely to have a lower asymmetry score based on the magenta and possibly the blue track being captured in the asymmetry calculation.

Global asymmetry score

The global asymmetry score is similar to the local asymmetry score but considers calorimetric contributions from all clusters in the event. A higher degree of global asymmetry can imply a more likely candidate as the combined momentum of the products (including invisible products) of the neutrino interaction should point close to the beamline direction to conserve momentum.

The global asymmetry score is calculated on a per-view basis and then resulting scores are then summed. The calculation of the local event axis is identical to the local asymmetry score (including the 5 cm search radius) but, following this, all the hits in the view are projected onto the axis and the asymmetry of their charge distribution about the candidate is calculated. In Figure 5.1, for instance, candidates near A, D and E are likely to be favoured based on the projections of all hits onto the local event axis, whereas candidates near B and C are likely to be disfavoured.

Shower asymmetry score

The shower asymmetry score considers all showers with at least one hit within 4 cm of the vertex candidate. The logic is identical to the local asymmetry score calculation except only the closest shower (if any) will be considered in each view. Consider Figure

5.3; vertices placed in the middle of the electron will receive a lower score if constituent 2D clusters are correctly identified as shower-like.

Event modifiers

We additionally consider a number of event-level features that are the same for every candidate in a given event but serve to modify the importance or interpretation of the candidate-level features. For example, events which are energetically sparse or dense may alter the relative importance of geometric features, and 'showeryness' and longitudinality alter the prior probabilities of different neutrino final states, which in turn affect the best vertex selection strategy. The event modifiers are:

- **Event charge:** The total charge deposited in the event across all views.
- **Event volume:** The volume of the smallest 3D bounding box that envelops all the clusters whose axes are aligned with the detector x, y and z axes.
- **Event longitudinality:** Considering the event volume bounding box, the ratio of the z length to the perimeter of the x-y rectangular cross-section: $\propto z/(x+y)$.
- **Event showeryness:** The proportion of hits associated with shower-like clusters across all views.
- **Number of hits:** The total number of 2D hits in the event across all views.
- **Number of clusters:** The total number of 2D clusters in the event across all views.

5.3 Model building

The features defined in the last section describe a 12-dimensional space that captures a variety of pieces of information about the event and the vertex candidate. To select the best candidate requires mapping points in this space to scalar scores. The features are not naturally probabilistic and exhibit nontrivial interdependence, so a simple weighted sum is unlikely to give the best solution, not to mention introducing at least 11 tunable constants. The task is suited to a supervised ML approach. Ranking a list of arbitrary

size is not immediately a classification or regression task. Machine-learned ranking (MLR) is itself an active field of research and MLR algorithms can be split into three categories:

- **Pointwise:** The problem is condensed into a regression problem; for each sample, we train a regressor to predict the final score. This assumes that good scores can be estimated for ground truth.
- **Pairwise:** The problem is condensed into a binary classification problem; for each pair of samples, we train a classifier to identify the better one. This assumes that good pairwise orderings can be estimated for ground truth.
- **Listwise:** We directly try to optimize a ranking evaluation metric such Spearman's rank correlation coefficient. This assumes that a good ranking can be estimated for ground truth.

In our case, we must consider what information is available to us as ground truth. We have the list of vertex candidates and the spatial position of the true vertex. Distance from the true vertex is a poor vertex quality metric for training, since a vertex at the wrong end of a muon, for example, is arguably higher quality than one in the middle of it, despite being further away from the true vertex. Scores calculated by previous algorithms are similarly not ideal if the goal is to perform significantly better. Instead, we take the N vertex candidates and label the one closest to the true vertex position as correct if it lies within 1 cm of the true vertex position. The $N - 1$ remaining candidates are labelled as incorrect. This is sufficient information to adopt a pairwise MLR approach: each event generates $N-1$ pairwise comparisons comprising the unique correct candidate versus each incorrect candidate.

The pairwise comparison can be performed using any binary classifier. We note that this was the first time that an ML algorithm was incorporated inside Pandora, a C++ project where development is encouraged to be dependency-light, highly performant and CPU-only. We therefore choose to use a simple SVM with a radial basis function kernel. We train external to Pandora using the Python library sklearn; inference inside Pandora is done using a custom C++ implementation ported from sklearn. In practice, we found that this setup tended to get the vertex roughly correct (i.e. within the correct

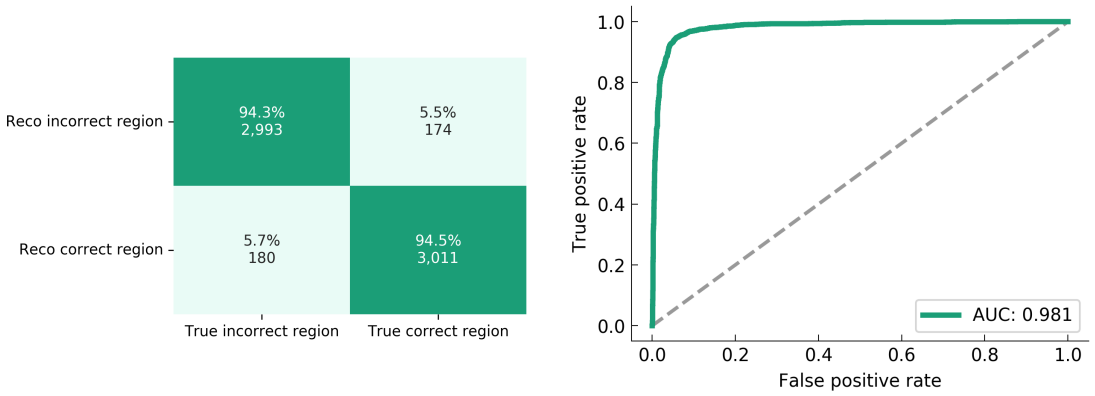


Figure 5.5: The performance of the region classifier: the confusion matrix (left) and ROC curve (right) across the simulated BNB ν_μ spectrum.

$\sim 2\,\text{cm}$ region) but otherwise performed worse than the `RPhi` algorithm at finding the single best vertex. In response to this, we split vertex selection into two distinct tasks:

1. **Region selection:** Optimized for selecting the best candidate from a set of candidates corresponding to the centres of non-overlapping circular regions of radius $5\,\text{cm}$.

2. **Single vertex selection:** Optimized for selecting the best single vertex candidate from a number of candidates with in a region of radius $2\,\text{cm}$.

We train a different SVM for each task. The intuition here is that region selection is more dependent on e.g. beam deweighting and asymmetry scores, whereas single vertex selection may depend more on r-ϕ and energy kick scores. Following a feature selection process, we drop the r-ϕ score from the region SVM.

5.3.1 Region model

The training data for the region model is constructed by first scoring every candidate using the `RPhi` algorithm, then using this list to create regions of vertices separated by at least $10\,\text{cm}$, using higher-ranked candidates to preferentially define and represent each region. Consider Figure 5.1; this view shows at least 12 vertex candidates but likely only four or five regions, depending on the distance between B and C. We train the SVM

classifier pairwise, as described in the previous section, where the region containing the true vertex position is taken as the correct region, if one exists.

SVM hyperparameters are optimized via k-fold cross-validation. The distribution of the candidate features used during training is shown in Figure 5.6; The distributions differ significantly for the correct region and the other regions. This is reflected in the confusion matrix[2] and receiver operating characteristic (ROC) curve shown in Figure 5.5. This simple feature model achieves an area under the ROC curve (AUC) of 0.981 and an overall accuracy of 94.4% compared with a baseline of 50.1%. Accuracy baselines in this chapter are defined by the classifier that always chooses the most likely class; due to class imbalance, this is not always 50%.

5.3.2 Vertex model

Let there be N vertex candidates within a 2 cm radius of the central hit of the true best region. The training data for the single vertex selection model comprises the $N-1$ pairwise comparisons between the candidate closest to the true vertex and each of the other candidates within a 2 cm radius of the central hit. Take again Figure 5.1; assuming E is chosen as the best region, it looks like at least 5 vertices (and therefore 4 pairwise comparisons) would be generated based on this region.

This model is trained entirely separately from the region model and SVM hyperparameters are optimized independently. The distribution of the candidate features used during training are given in Appendix A. The distributions here are much more similar but vary in the expected way: the true vertex tends to have higher beam deweighting, asymmetry and r-ϕ scores and a lower energy kick score. The similarity of these distributions is apparent in the performance, with an AUC of 0.699. The overall accuracy is 62.7% compared with a baseline of 50.6%. Full performance details can similarly be found in Appendix A. Based on these features, picking the true best vertex from generated candidates with 2 cm of each other appears more challenging that choosing the best region, though this is a function of the similarity and density of candidates within

[2]The columns of a confusion matrix represent the true labels while the rows represent the (equivalently ordered) predicted labels. The diagonal elements therefore represent correct predictions and the off-diagonal elements incorrect predictions. The percentages are produced by normalizing by each column (the true label).

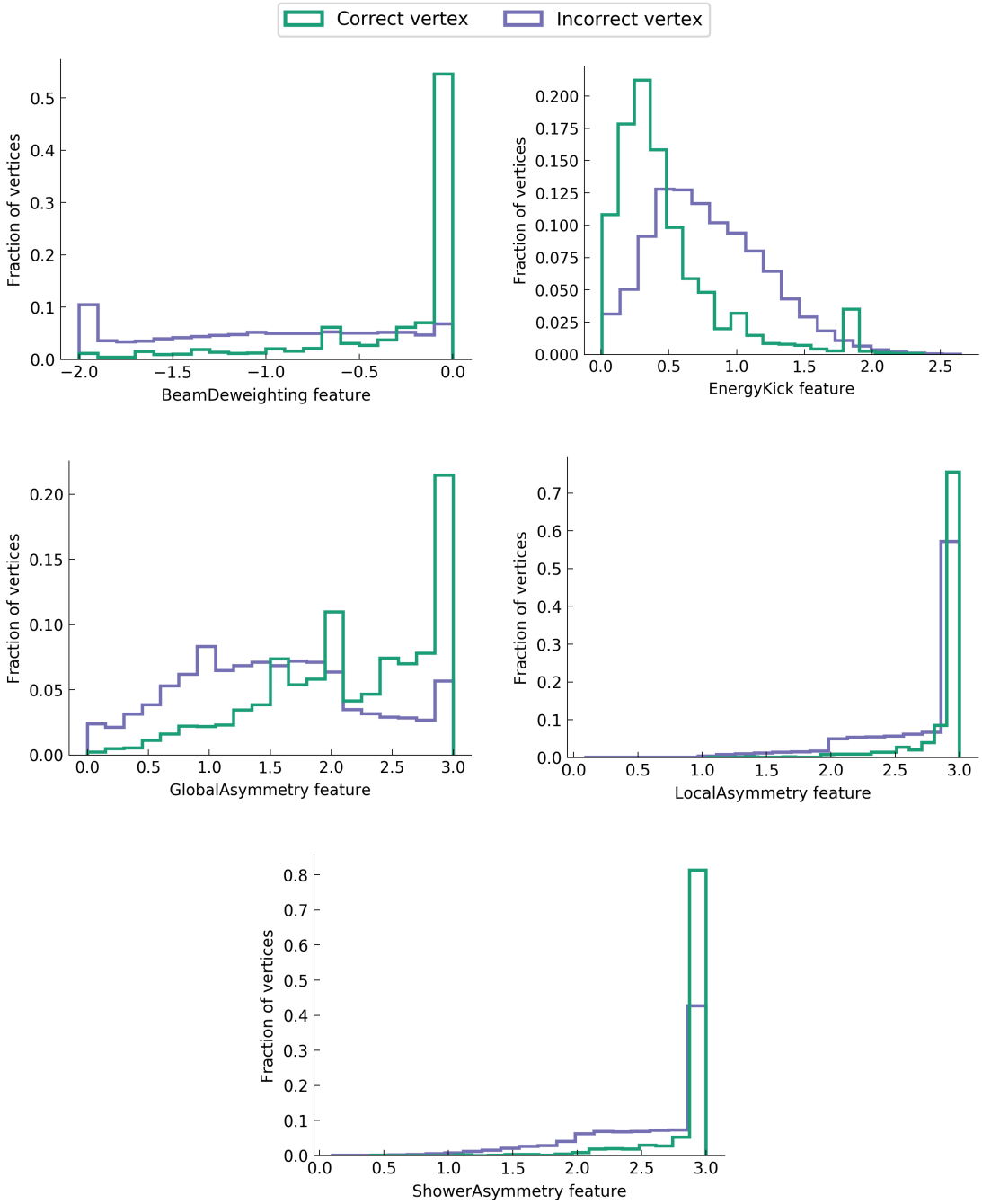


Figure 5.6: The distributions of features used in the region selection algorithm for the correct and incorrect region candidates across the simulated BNB ν_μ spectrum.

Table 5.1: Comparison of the SVM and `RPhi` vertex selection algorithms. We consider the performance for simulated BNB ν_μ events, simulated BNB intrinsic ν_e events and a small number of selected channels from each dataset. We use the proportion of events with a reconstructed neutrino interaction vertex within 5 cm of the true vertex position ($\Delta_{\text{RPhi}} < 5\,\text{cm}$) to characterize performance.

Simulation	Channel	$\Delta R_{\text{RPhi}} < 5\,\text{cm}$ (%)	$\Delta R_{\text{SVM}} < 5\,\text{cm}$ (%)
BNB ν_μ	All	59.2	**63.9**
	`CCQEL_MU`	**49.0**	45.9
	`CCQEL_MU_P`	80.8	**85.2**
	`CCRES_MU_PIZERO`	56.3	**71.2**
BNB intrinsic ν_e	All	53.6	**59.3**
	`CCQEL_E`	**47.0**	39.0
	`CCQEL_E_P`	70.4	**74.6**
	`CCRES_E_PIZERO`	45.1	**60.9**

the region.

5.4 Results and discussion

Classifier performance is a proxy for actual improvements to the vertex reconstruction quality, which we characterize using the distribution of the absolute error in the spatial position of the chosen vertex: Δx, Δy, Δz and $\Delta R = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$. We wish for these errors to be as peaky and close to 0 as possible, and we compare them with those of the `RPhi` algorithm. We also assess the proportion of events that achieve $\Delta R < 5\,\text{cm}$ and the fraction of correctly reconstructed events (as per the definition in 4), both as a function of true neutrino energy.

The spatial error metrics, fraction of good vertices and fraction of correct events are shown in Figure 5.7 for simulated BNB ν_μ events and in Figure 5.8 for simulated BNB intrinsic ν_e events, both considering all channels. A summary of key results for some selected channels is given in Table 5.1. For simulated BNB ν_μ events, the proportion of vertices within 5 cm of the true vertex increases from 59.2% to 63.9%; for intrinsic ν_e, it increases from 53.6% to 59.3%. This benefit is consistent across a broad energy spectrum and seems to address the drop-off in performance of the `RPhi` algorithm at higher energies. The fraction of correct events improves slightly at low energies but

otherwise does not change significantly, suggesting other error modes may be dominant when considering events across all channels.

We note an improvement across almost all channels, including significant improvement for more complex topologies such as CC resonant $\mu + \pi^0$ interactions, and a slight decrease in performance for single muons and single electrons. Despite overall improved performance, channels involving muons (μ, $\mu + p$, $\mu + \pi^0$) experience a slight drop-off at high energies compared with the `RPhi` algorithm. These are due to the trade-off introduced by the introduction of richer information; the simpler `RPhi` algorithm relies heavily on the beam deweighting factor, boosting its performance for single particles and events with long tracks but making it less suitable for more complex topologies. For the SVM algorithm, δ-rays in the middle of muons can lead to misplaced vertices, an effect which is more likely for longer muons, hence the drop-off at higher energies. Likewise, single electrons often begin track-like before producing electromagnetic showers (see Figure 5.3), which can lead to two distinct clusters, one tagged as shower-like and one as track-like. This can lead the SVM algorithm to consider the point where the clusters meet as an appealing vertex position; the asymmetry scores do not protect against this as it is in the middle of neither a track nor a shower.

More detailed results, including feature distributions and performance metrics on a per-channel basis for selected channels, can be found in Appendix A. The quality of vertex reconstruction affects analysis not just through reconstruction correctness but also through direct effects on energy reconstruction due to effects on particle identification and range estimation. This will be studied in Chapter 9.

This has been the default vertex reconstruction algorithm at MicroBooNE since 2017. The dramatic improvement in the reconstruction of complex topologies has enabled new analyses with up to six visible particles in the final state [127]. Particle identification and especially electron/photon discrimination, crucial to MicroBooNE's physics goals (see Section 3.1.2), require fine-resolution vertex reconstruction to preserve event topology and enable high-quality calorimetry. The improvements presented in this chapter directly contribute to this goal, as well as representing a proof-of-concept for applying machine-learned ranking to neutrino interaction vertex reconstruction. Future work in this vein, combining more expressive deep learning methods with better topo-

logical/calorimetric representations of the event and thoughtful inductive biases, will likely perform even better.

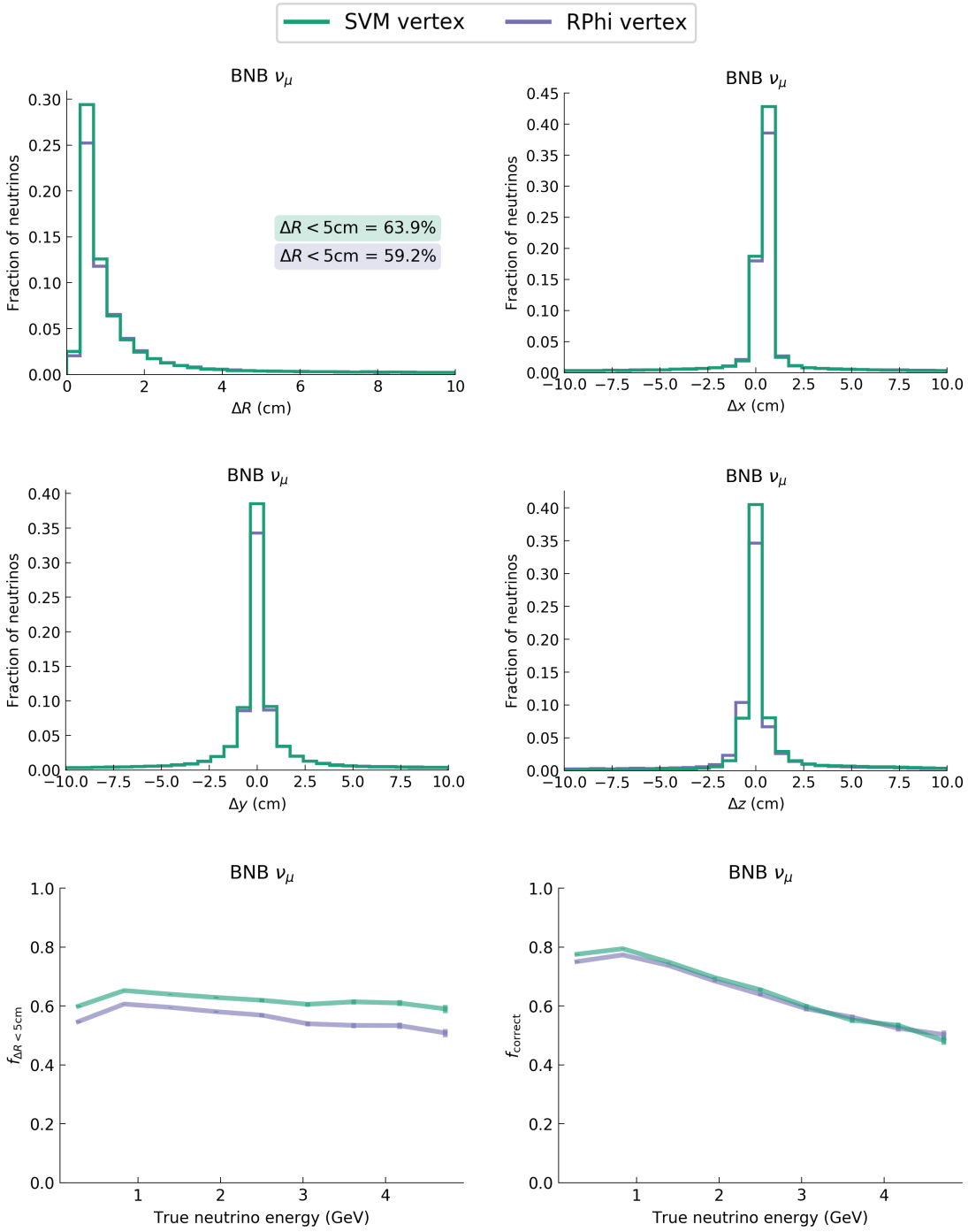


Figure 5.7: Vertex reconstruction quality for the two algorithms across the simulated BNB ν_μ spectrum. The top two rows show the effect of the new algorithm on Δx, Δy, Δz and ΔR compared with the true vertex. The bottom left and right plots show the fraction of vertices within 5 cm of the true vertex and the fraction of correct events, respectively.

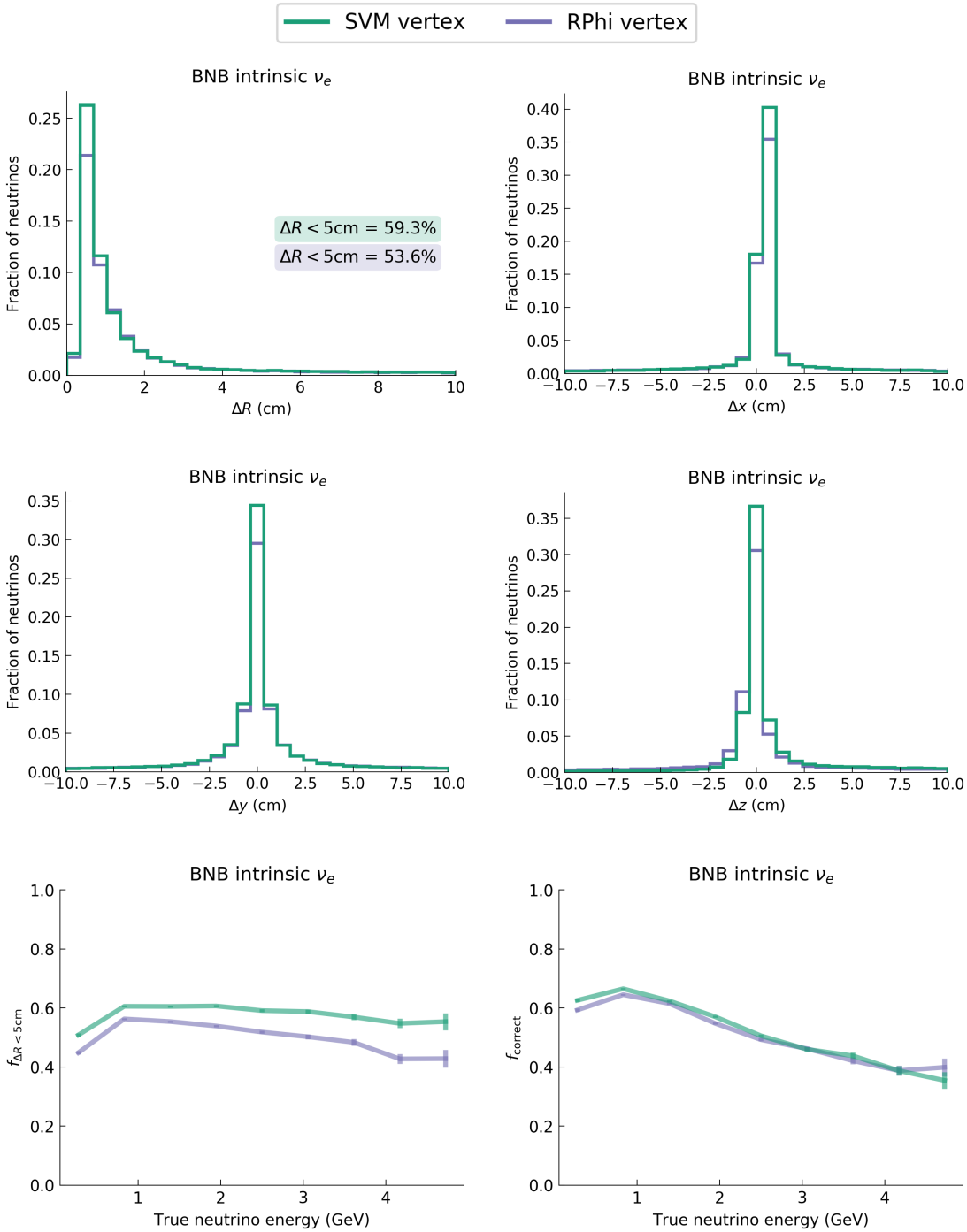


Figure 5.8: Vertex reconstruction quality for the two algorithms across the simulated BNB intrinsic ν_e spectrum. The top two rows show the effect of the new algorithm on Δx, Δy, Δz and ΔR compared with the true vertex. The bottom left plot shows the fraction of vertices within 5 cm of the true vertex and the bottom right plot the fraction of correct events, both as a function of true neutrino energy.

Chapter 6

Modelling $\mathrm{d}Q/\mathrm{d}x$ at low energies

This chapter introduces the theory of energy losses by charged particles traversing detectors and derives mathematical models for $\mathrm{d}E/\mathrm{d}x$ and $\mathrm{d}Q/\mathrm{d}x$ at low energies, i.e. $T \ll mc^2$. Two models of electron-ion recombination, Birks' law and the modified box model, are compared and incorporated into the approximation. Two novel, analytic approximations of modal $\mathrm{d}Q/\mathrm{d}x$ models are derived, with potential application to particle identification.

6.1 Introduction

As charged particles traverse a detector medium, they lose energy. The ability to model this effect as a function of the distance traversed by a particle is critical both to particle identification and energy estimation, of crucial importance to any experiment. A statistical distribution for such energy losses was first described by Landau in 1944 [128], and later generalized by Vavilov [129] in 1957. These distributions, after the inclusion of modern corrections, still form the foundations of our understanding of energy losses by charged particles in detectors.

In a liquid argon time-projection chamber (LArTPC) such as the MicroBooNE detector, the kinetic energy of the incident particles is primarily transferred to bound electrons in the medium, leading to an ionization trail observed as a channel- and time-dependent current. This process is subject to various effects, including electron diffusion, electron attenuation, and hardware issues such as quiet or noisy channels. After undoing

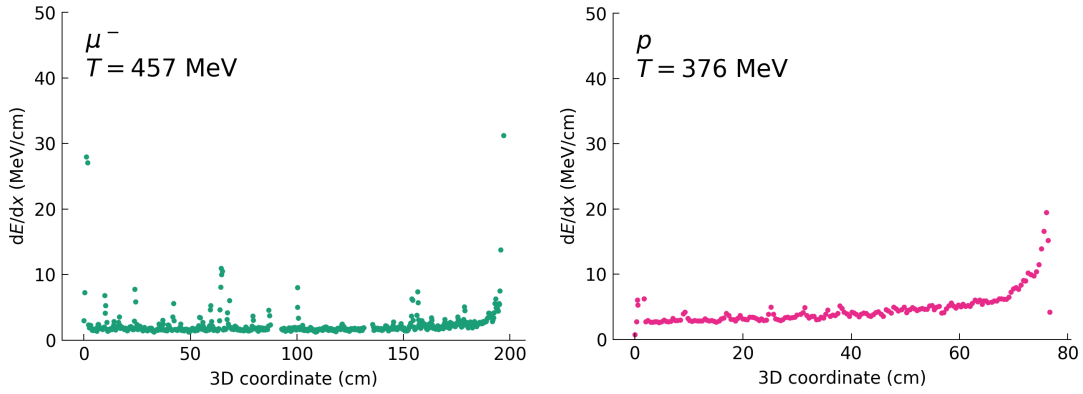


Figure 6.1: Reconstructed $\mathrm{d}E/\mathrm{d}x$ as a function of 3D track coordinate for a simulated muon (left) and a simulated proton (right) at MicroBooNE.

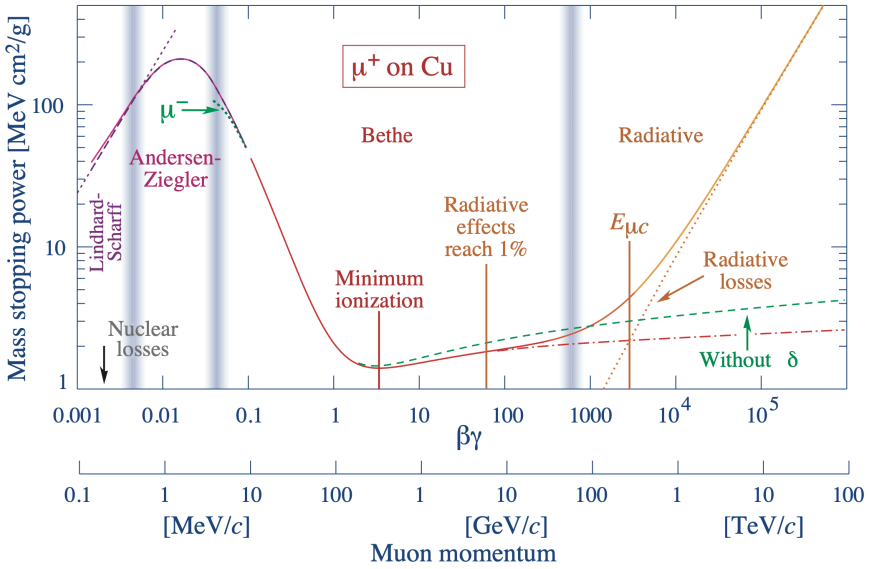


Figure 6.2: Mass stopping power (proportional to $\mathrm{d}E/\mathrm{d}x$ in this chapter) for positive muons in copper as a function of $\beta\gamma = p/Mc$. Domains of applicability of different approximations is indicated by vertical bands. The solid curves indicate the overall mass stopping power. Figure from [13].

these effects to the extent possible, we are left with charge depositions in discretized 'hits' with widths and positions in the appropriate wire plane, as described in Chapter 3. After reconstruction, it is then possible to infer approximate 3D locations and extents of these depositions (as demonstrated in Chapter 7), thus enabling estimation of the behaviour of $\mathrm{d}Q/\mathrm{d}x$ along the track, which can then be transformed to form an

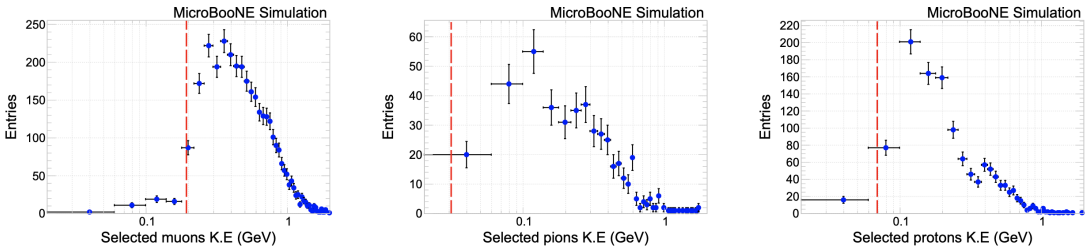


Figure 6.3: Simulated energy distributions for muons, charged pions and protons from a MicroBooNE study of charged particle multiplicity. This includes only particles which passed an event selection designed to select high-quality tracks and reduce cosmic background. The red dotted lines indicate a particle range cut later used in that study. Figure from [127].

estimate of $\mathrm{d}E/\mathrm{d}x$. Examples of observed $\mathrm{d}E/\mathrm{d}x$ along the tracks of a simulation muon and proton in MicroBooNE are shown in Figure 6.1.

The character of this observed ionization pattern is a function of both particle and detector properties and, as such, different particles have different calorimetric signatures in a given detector. The observed energy losses and their relationship with the distance traversed by the particle can therefore be exploited for particle identification and energy estimation. Our approach in this chapter is to model mean/modal energy losses with a view to aiding particle identification. A particle identification technique based on this work is presented in Chapter 7. A complementary deep learning approach is developed in Chapter 8.

We restrict our discussion to electronic and radiative energy losses by heavy track-like particles of moderate energies passing through matter. "Moderate" means $0.1 \lesssim \beta\gamma \lesssim 1000$ for the particles we will consider, such that mean losses are adequately described by the Bethe equation (see Figure 6.2. For muons, this amounts to kinetic energy in the range $10\,\mathrm{MeV} \lesssim T \lesssim 100\,\mathrm{GeV}$; for protons, $100\,\mathrm{MeV} \lesssim T \lesssim 1000\,\mathrm{GeV}$. In particular, we will neglect nuclear losses (dominant at energies below the MicroBooNE detection threshold) and large radiative losses (dominant at very high energies). Energy distributions for muons, charged pions and protons from a charge multiplicity study are shown in Figure 6.3; this condition is well-satisfied for the majority of particles at MicroBooNE. With application to LArTPCs, we will focus on modelling energy losses by muons, protons, charged pions and charged kaons through ionization, and include a

discussion of radiative losses by muons.

6.2 Energy losses by charged particles

6.2.1 The Bethe formula

Fast-moving charged particles in a moderate energy range primarily lose their energy through single collisions with electrons in the medium [13]. For detectors of sufficient thickness (including LArTPCs), such collisions are frequent but exchange relatively small amounts of energy, permitting accurate description by continuous statistical distributions. For a particle of mass M, the maximum transfer of energy possible in a single collision can be shown to be [13]

$$W_{\max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2}. \tag{6.1}$$

By applying a correction to the Rutherford cross-section, which describes collisions with free electrons, Bethe derived a model for collisions with electrons bound in nuclei. The 'Bethe formula' accurately models the *stopping power* (i.e. the mean rate of energy loss) of charged particles in the mass-energy region $0.1 \lesssim \beta\gamma \lesssim 1000$ traversing media with intermediate values of atomic number Z [130]:

$$-\left\langle \frac{\mathrm{d}E}{\mathrm{d}x} \right\rangle = \rho K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right], \tag{6.2}$$

where

- $-\mathrm{d}E/\mathrm{d}x$ is the rate of energy loss of the particle per unit distance traversed[1] in MeV/cm;
- ρ is the density of the medium in g/cm^3;
- $K = 4\pi N_A r_e^2 m_e c^2 \approx 0.307075\,\mathrm{MeV\,mol^{-1}\,cm^2}$ is a constant (where N_A is the Avogadro constant and r_e the classical electron radius);
- Z is the atomic number of the absorber;

[1]Note that the Bethe formula here differs from its statement in [13] by a factor of ρ due to our preferred units, i.e. preferring the 'linear stopping power' to the 'mass stopping power'.

- A is the atomic mass of the absorber in $\mathrm{g\,mol^{-1}}$;
- z is the charge number of the incident particle;
- I is the mean excitation energy of the absorber in e.g. MeV, to match the units of $W_{\max}$ (as defined in Equation 6.1) and the electron mass $m_e c^2$; and,
- $\delta(\beta\gamma)$ is the density effect correction term[2], whose meaning and form will be discussed in 6.2.5.

An example of the Bethe mean energy loss is plotted in Figure 6.4 for a 1 GeV muon in a LArTPC, along with the kinetic energy profile of the particle. This figure, along with all other calorimetric simulations in this chapter, was produced using our C++ library `bethe-faster` [132], a framework for particle identification and energy estimation using a sequential Monte Carlo method. A few important features may be noticed: a roughly constant amount of ionization at high energies, a sharp rise at the end of the track (known as the Bragg peak), and a minimum preceding the sharp rise. These features are common to all charged particles in similar conditions, though their magnitude is a function of particle and detector parameters. The dependence on particle mass, in this case, is the key to performing particle identification using calorimetric data. The minimum mean energy loss attained by a muon can be seen to be approximately 2.2 MeV/cm. A particle travelling with this energy loss is called a *minimum ionizing particle* (MIP).

For tracks at MicroBooNE, we observe these kinds of calorimetric signatures but superposed with correlated high-$\mathrm{d}E/\mathrm{d}x$ noise to the distribution, as per Figure 6.1, which shows the simulated $\mathrm{d}E/\mathrm{d}x$ profile of a muon and a proton along their track length. The origin of some of this noise is the shape of the Landau-Vavilov distribution, whilst some is the result of small δ-rays: secondary electrons emitted by the incident particle whose charge deposition was subsumed by the parent particle's hits. This renders the measured charge too high compared with the 3D path measurement. For the MC simulation, it is possible to remove this effect by weighting each charge deposition by its proportion induced by the desired MC particle.

[2]This is a more modern (1952) addition by Sternheimer, not originally included by Bethe; see [131].

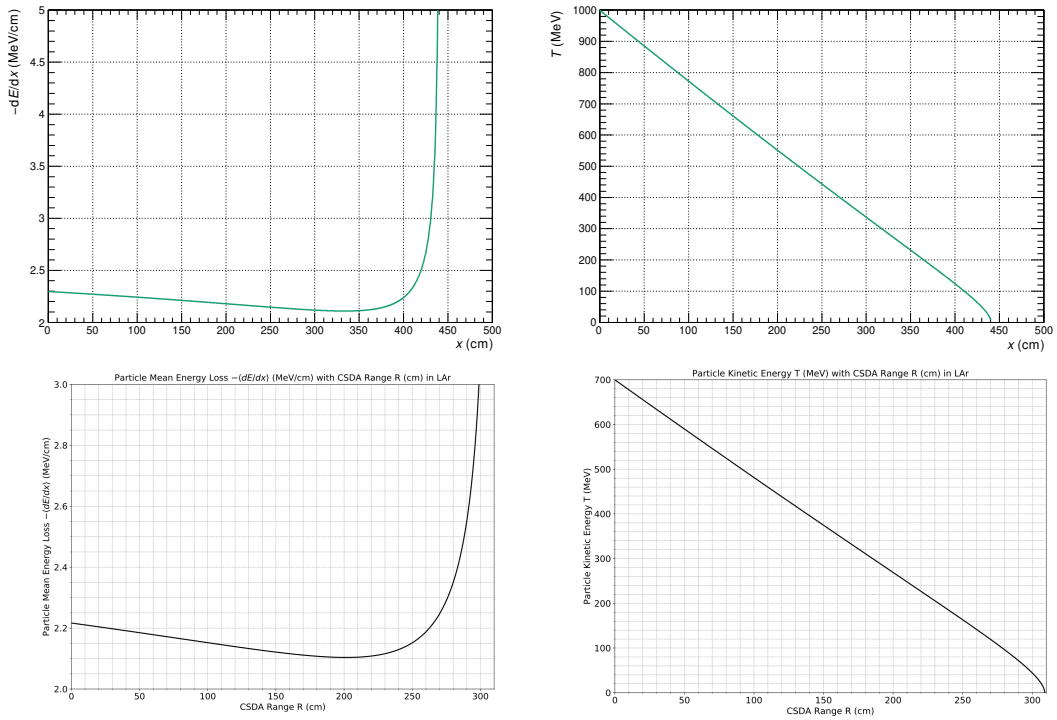


Figure 6.4: Simulated evolution of the Bethe mean energy loss (top left) and resulting kinetic energy (top right) of a muon in a LArTPC with initial kinetic energy of 1 GeV. Note that the particle is only rendered in the Bethe range of applicability $0.1 \lesssim \beta\gamma \lesssim 1000$. The kinetic energy is calculated under the continuous slowing-down approximation (see main text) and the numerical integration is performed using `bethe-faster` [132]. Bottom-row plots are the equivalent plots from an independent study [109].

6.2.2 Continuous slowing-down approximation

In order to model the kinetic energy of the particle or the energy deposition of a particle in a detector using the Bethe formula, it must be integrated:

$$\Delta E_{\text{Bethe}}(x_0 \to x_1) = \\ -\rho K z^2 \frac{Z}{A} \int_{x_0}^{x_1} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \mathrm{d}x, \quad (6.3)$$

noting that β, γ and δ all have a nontrivial dependence on x. This equation permits no analytical solution[3] but may be evaluated numerically. Numerical integration of the Bethe formula was used to produce the plot of kinetic energy in Figure 6.4, in which we also compare equivalent plots from an independent study. Both top and bottom row plots agree that the range of a 700 MeV muon is $\sim$ 310 cm, giving confidence in the

[3]This is true whether we use the Sternheimer parametrization for the density effect correction term or omit it entirely.

method presented. These calculations (usually stored a lookup tables) are extensively used in particle physics for calculating particle energy from range.

In reality, the energy lost by the particle over a given range follows a statistical distribution, of which Equation 6.3 describes the mean. Using this equation to model the dynamics of the particle is known as the *continuous slowing-down approximation* (CSDA) [133]: we neglect fluctuations and assign the same energy loss to every point in the track. In practice, this can provide a reasonable estimate of a stopping particle's range, as will be demonstrated in Section 6.2.3.

6.2.3 Most probable energy loss

In [13], the authors complain of the misuse of the Bethe equation in high-energy physics. The basis of their complaint is that the energy loss distribution is skewed by rare but often very large losses, such that the mean loss is significantly higher than the most probable loss. Observations of single events (or even hundreds of events) can, then, be better-described by the most probable loss.

An expression for the most probable energy loss for detectors of thickness l in a moderate range can be derived from the Landau-Vavilov model [129]:

$$-\Delta E_{\mathrm{mp}} = \xi \left[\ln \frac{2m_e c^2 \beta^2 \gamma^2}{I} + \ln \frac{\xi}{I} + j - \beta^2 - \delta(\beta\gamma) \right] \tag{6.4}$$

where

$$\xi := z^2 \frac{K}{2} \left\langle \frac{Z}{A} \right\rangle \frac{l\rho}{\beta^2}, \tag{6.5}$$

l is the detector thickness[4] in cm, and j is a constant of value ~ 0.2000. All other parameters are as described for Equation 6.2. Whilst the mean energy loss is independent of detector thickness l, the modal energy loss in an increasing function of l.

It should also be noted that, in the high-energy limit, the behaviour of the density effect correction term $\delta(\beta\gamma)$ is such that $-\Delta E_{\mathrm{mp}}$ reaches a *Fermi plateau* [131]:

$$-\Delta E_{\mathrm{mp}} \xrightarrow[\beta\gamma \gtrsim 100]{} -\Delta E_{\mathrm{Fermi}} = \xi \left[\ln \frac{2m_e c^2 \xi}{(\hbar\omega_p)^2} + j \right], \tag{6.6}$$

[4]The statement of ξ here differs from its statement in [13] by a factor of ρ since we express l in units of cm instead of $\mathrm{g\,cm^{-2}}$.

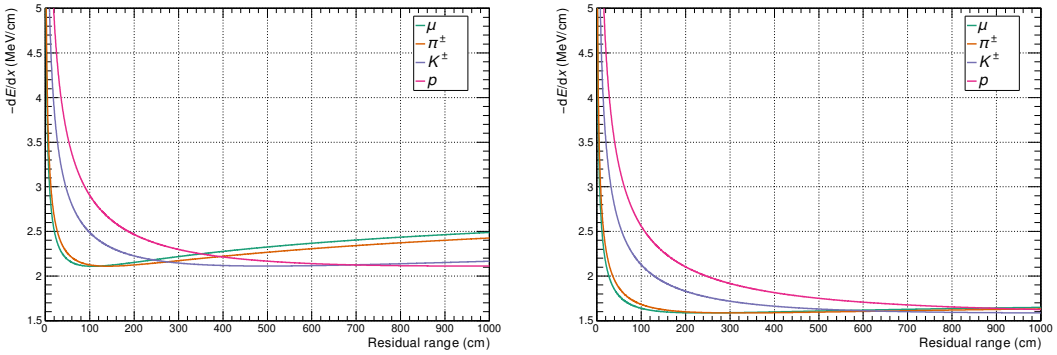


Figure 6.5: Simulated evolution of the linear stopping power for muons, charged pions, charged kaons and protons, using the mean stopping power (left) or the modal stopping power (right) as a function of the residual range. Note that the modal stopping power is significantly lower than the mean stopping power, including at minimum ionization.

since

$$\delta \xrightarrow[\beta\gamma \gtrsim 100]{} \delta_{\text{Fermi}} = 2 \ln \frac{\hbar\omega_p \beta\gamma}{I} - 1, \tag{6.7}$$

where

$$\hbar\omega_p := \frac{m_e c^2}{\alpha} \sqrt{4\pi N_e r_e^3} \approx \sqrt{\rho \left\langle \frac{Z}{A} \right\rangle} \times 28.816\,\text{eV} \tag{6.8}$$

is the plasma energy [13], α is the fine-structure constant $e^2/4\pi\epsilon_0\hbar c$, and N_e is the electron density of the medium. The canonical parametrization of the density effect correction will be given in Section 6.2.5.

Figure 6.5 demonstrates the significant difference between the mean and modal stopping power for a variety of particles in MicroBooNE as a function of the *residual* range (i.e. the coordinate subtracted from the maximum range). While the minimum mean ionization of a muon is at $\sim 2.1\,\text{MeV/cm}$, the minimum modal ionization of a muon is at the much lower value of $\sim 1.4\,\text{MeV/cm}$. The plots show the stopping power as a function of the residual range rather than the coordinate itself to facilitate comparison between different particles. Figure 6.6 similarly shows the difference between mean and modal stopping power against the backdrop of a simulated LArTPC energy loss distribution. The modal stopping power predicts a much longer range than both the mean and stochastic stopping powers due to the skew of the distribution. The simu-

lated stochastic stopping power takes into account other effects that truncate very large energy losses, so tends to have a larger range than the mean (CSDA) estimate.

Further, at high energies far from the Bragg peak, the modal loss (unlike the mean loss) provides little discriminatory power between particle masses. The modal loss as a function of kinetic energy is shown in Figure 6.7. For non-stopping particles[5], we may never reach low enough energies to observe a Bragg peak and so particle identification based on calorimetric information alone could be difficult or infeasible. Note that the modal calorimetric behaviours of muons and pions are almost indistinguishable, even in this idealized setting.

We wish to use observations derived from the energy loss distributions of single particles to perform particle identification, which may only comprise a handful of data points. Since estimating the mean stopping power reliably requires a large amount of data, a method based on the estimated mode will likely be more fruitful. In particular, the behaviour of the estimated mode at low energies has particular discriminatory power, as will be demonstrated.

6.2.4 Dependence of most probable energy loss on l

Equation 6.2 suggests that the mean energy loss $-\langle \mathrm{d}E/\mathrm{d}x \rangle$ is independent of the cell size l. This must be the case to ensure that the expected loss over two layers of width L is the same as the expected loss over a single layer of width $2L$; i.e. the length at which we effectively sample the particle should not affect the rate at which it loses energy. However, the expression for the modal loss given in Equation 6.4 suggests that $-\Delta E_{\mathrm{mp}}/l \propto a \ln l + b$; i.e. the modal loss over a layer of width $2L$ is larger than the modal loss over two layers of width L. This somewhat unintuitive result is a sampling effect present in both theory and practice [13, 134]. This is illustrated in Figure 6.8.

This sampling effect does not mean that the actual behaviour of a particle changes based on the cell width. A measurement of the modal $\mathrm{d}E/\mathrm{d}x$ along a track *is* dependent on l (which is a function of the track direction) but a measurement of the total ΔE is unchanged. A plot of modal $\mathrm{d}E/\mathrm{d}x$ does not represent real losses (see Figure 6.6) but

[5] I.e. those that decay in-detector, leave the detector or interact in a way other than to produce ionization electrons before reaching an energy low enough for us to observe a significant Bragg peak.

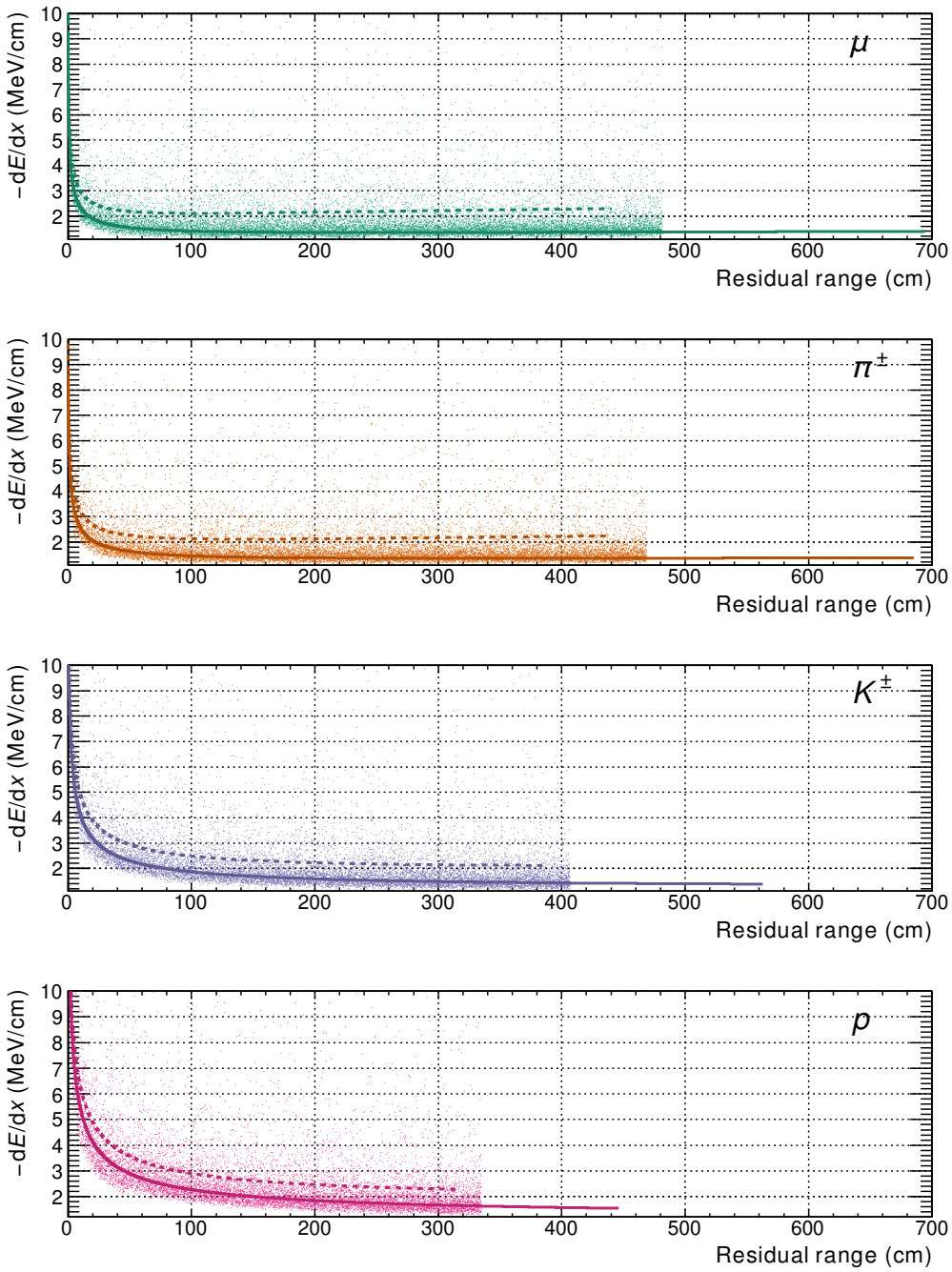


Figure 6.6: Comparison of the simulated modal (solid), mean (dashed) and stochastic (points) energy losses for 1 GeV muons, charged pions, charged kaons and protons in a MicroBooNE-like detector. Note that the modal loss well-characterizes the observed distribution but provides a poor range estimate in this case; the range implied by the mean loss (corresponding to the CSDA) is closer to the truth.

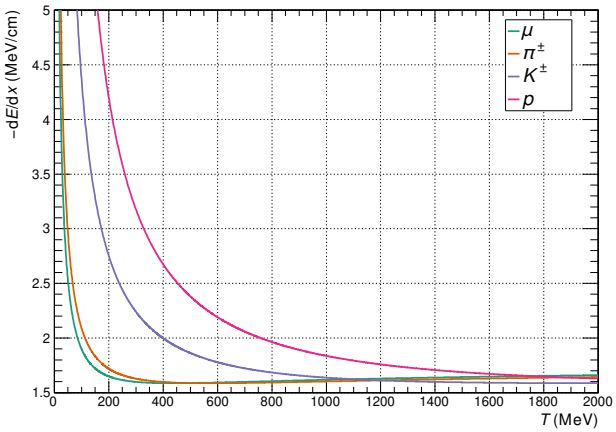


Figure 6.7: Simulated evolution of the modal linear stopping power for muons, charged pions, charged kaons and protons as a function of the instantaneous kinetic energy of the particle. Made using `bethe-faster` [132].

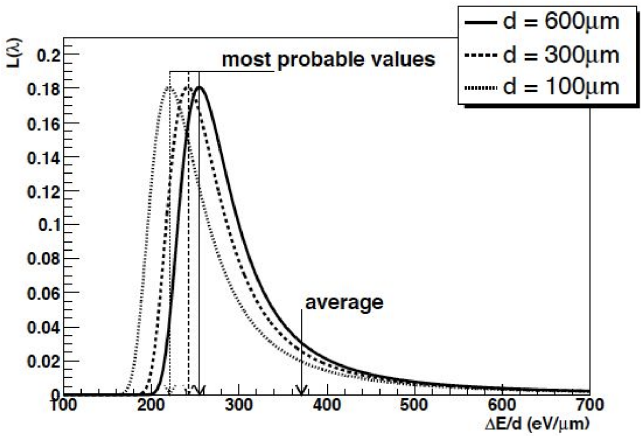


Figure 6.8: Distribution of $\Delta E/l$ (here denoted $\Delta E/d$) for a MIP in silicon with a thickness d, following the Landau distribution. The dependence of the mode on d but the invariance of the mean is indicated. Figure from [134].

can be used as a distribution characterization tool to perform particle ID.

6.2.5 Density effect correction

The functional form used for the density effect correction δ is Sternheimer's parametrisation [131]:

$$\delta_{\text{Stern}} = \begin{cases} 2\,(\ln 10)\,s - \bar{C} & \text{if } s \geq s_1 \\ 2\,(\ln 10)\,s - \bar{C} + a(s_1 - s)^k & \text{if } s_0 \leq s < s_1 \\ 0 & \text{if } s < s_0 \text{ (nonconductors)} \\ \delta_0 10^{2(s-s_0)} & \text{if } s < s_0 \text{ (conductors)} \end{cases} \tag{6.9}$$

where $s := \log_{10}(p/mc) = \log_{10}(\beta\gamma)$, and $\bar{C}$ is chosen to correspond to the high-energy ($s \geq s_1$) limit in 6.7, *viz.*

$$\bar{C} = 2 \ln \frac{I}{\hbar\omega_p} + 1. \tag{6.10}$$

The dimensionless parameters s_0, s_1, a and δ_0 are to be derived experimentally for different media. A reference for these values is given in [135]; for liquid argon, the authors give best values of

- $a = 0.19559$,
- $k = 3.0000$,
- $s_0 = 0.2000$,
- $s_1 = 3.0000$, and
- $\delta_0 = 0.00$ (since liquid argon is not a conductor).

Note that, for liquid argon, the kinetic energy below which Sternheimer's δ is 0 for a particle of mass m is $\sim 0.87mc^2$ ($\sim 90\,\text{MeV}$ for a muon or $\sim 820\,\text{MeV}$ for a proton). We note that protons at MicroBooNE are often below this threshold (Figure 6.3). The kinetic energy above which Sternheimer's δ has reached the Fermi plateau is $999mc^2$ ($\sim 105\,\text{GeV}$ for a muon or $\sim 940\,\text{GeV}$ for a proton, far above MicroBooNE energies).

In practice, the density effect correction is usually negligible for $T < mc^2$ and appears roughly linear long before we reach the Fermi plateau. The relative magnitude $R_0(T)$

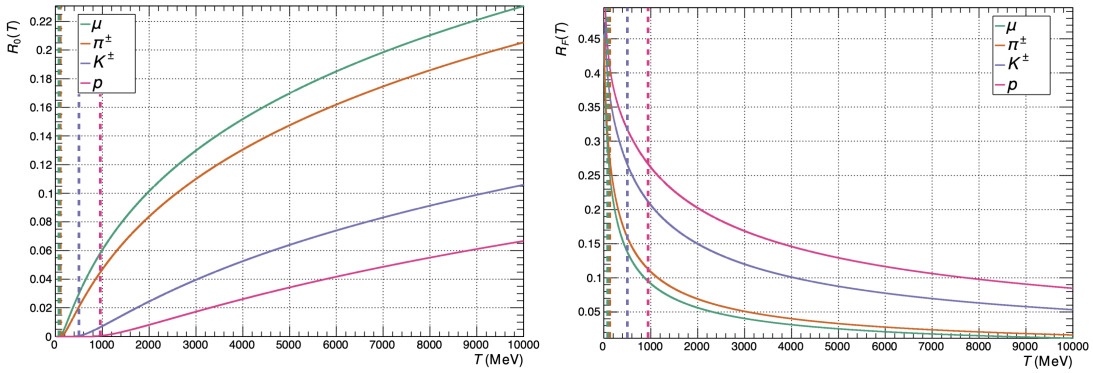


Figure 6.9: The fractional error in the modal $-\mathrm{d}E/\mathrm{d}x$ from assuming no density effect correction (left) or the high-energy limit of the density effect correction (right). The error is shown as a function of the instantaneous kinetic energy for a set of particles in a LArTPC. Neglecting density effect corrections in the regime $T \leq mc^2$ would yield an error $\lesssim 1\%$ for all particles. $T = mc^2$ is indicated with a dashed line for each particle.

of the density effect correction compared with other terms in Equation 6.4 is shown as a function of kinetic energy in Figure 6.9:

$$R_0(T) := \left| \frac{\delta(\beta\gamma)}{\ln(2m_e c^2\beta^2\gamma^2/I) + \ln(\xi/I) + j - \beta^2 - \delta(\beta\gamma)} \right|. \tag{6.11}$$

This is the fractional error incurred by ignoring density corrections for a given particle as a function of kinetic energy. We may similarly derive from Equations 6.4 and 6.6 the fractional error $R_F(T)$ incurred by assuming the Fermi plateau:

$$R_F(T) := \left| \frac{\delta(\beta\gamma) - 2\ln(\beta\gamma\hbar\omega_p/I) + \beta^2}{\ln(2m_e c^2\beta^2\gamma^2/I) + \ln(\xi/I) + j - \beta^2 - \delta(\beta\gamma)} \right|. \tag{6.12}$$

This is also shown in Figure 6.9. We will use these fractional errors to inform good approximations to Equation 6.4 in the low-energy limit.

6.3 Modelling $\mathrm{d}E/\mathrm{d}x$

6.3.1 Modal behaviour

The response of the detector to energy deposition is subject to various detector effects (including electron-ion recombination), so we will first try to develop a model for $-\mathrm{d}E/\mathrm{d}x$ from the theory described in Section 6.2, before using it to model the observed $\mathrm{d}Q/\mathrm{d}x$. We wish to discover whether a fast, analytic function exists of adequate quality to model particles in MicroBooNE in a useful regime of energy for particle identification.

Starting from Equation 6.4, we identify the most probable energy loss over a cell of thickness l:

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\mathrm{mp}} \sim \frac{-\Delta E_{\mathrm{mp}}}{l} \tag{6.13}$$

as per [13]. Understanding $-\mathrm{d}E/\mathrm{d}x$ to mean our approximation to the most probable loss, we then have

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{\xi}{l}\left[\ln\frac{2m_e c^2\beta^2\gamma^2}{I} + \ln\frac{\xi}{I} + j - \beta^2 - \delta(\beta\gamma)\right]. \tag{6.14}$$

We relate β^2 to the kinetic energy T for a particle of mass m via

$$T = (\gamma - 1)mc^2 = \left(\frac{1}{\sqrt{1-\beta^2}} - 1\right)mc^2 \tag{6.15}$$

$$\therefore \beta^2 = 1 - \frac{1}{\left(\frac{T}{mc^2}+1\right)^2}, \tag{6.16}$$

and, using 6.15, γ to T via

$$\gamma = 1 + \frac{T}{mc^2}. \tag{6.17}$$

To make the dependence of the equation on β^2 and l explicit, we also define

$$\xi' := \frac{\beta^2}{l}\xi = z^2\frac{K}{2}\left\langle\frac{Z}{A}\right\rangle\rho \tag{6.18}$$

from 6.5. We note that ξ' is a detector constant and has value $\sim 0.0969\,\mathrm{MeV/cm}$ for MicroBooNE assuming:

- liquid argon density $\rho = 1.40\,\mathrm{g\,cm^{-3}}$;
- atomic number of argon $Z = 18$;
- charge of incident particle $z = 1$; and,
- atomic mass of argon $A = 39.95\,\mathrm{g\,mol^{-1}}$.

Using[6] $-\mathrm{d}E/\mathrm{d}x = -\mathrm{d}T/\mathrm{d}x$, we may substitute 6.16, 6.17 and 6.18 to rewrite 6.14 as the first-order nonlinear ordinary differential equation

$$\left(1 - \frac{1}{\left(\frac{T}{mc^2}+1\right)^2}\right)\left(1 - \frac{1}{\xi'}\frac{\mathrm{d}T}{\mathrm{d}x}\right) - 2\ln\left(1 + \frac{T}{mc^2}\right) + \delta(T) - \ln\frac{2m_e c^2\xi' l}{I^2} - j = 0, \tag{6.19}$$

[6]The energy and the kinetic energy differ only by a constant offset: the rest mass.

where we are now expressing the density correction term δ as a function of T but note that it is also parametrized by the particle mass m. A solution to this equation $T(x; m, T_0)$ describes the modal energy estimate as a function of x for a particle of mass m with $T(x = 0) = T_0$, which could then be used in conjunction with the ODE to model modal $-\mathrm{d}E/\mathrm{d}x$ along a track. The modal T value used here hugely underestimates the true kinetic energy; this is evidenced by the implied range measurement using the solid lines in Figure 6.6. Instead, the mode is is a tool for characterizing the distribution for particle ID purposes. Whether using the Sternheimer parametrization for δ or neglecting it altogether, Equation 6.19 is not analytically soluble, though it is feasible to find numerical solutions.

For brevity, we introduce the unitless, scaled, modal kinetic energy $T' := T/mc^2$ such that Equation 6.19 becomes

$$\left(1 - \frac{1}{(T'+1)^2}\right)\left(1 - \frac{mc^2}{\xi'}\frac{\mathrm{d}T'}{\mathrm{d}x}\right) - 2\ln\left(1+T'\right) + \delta(T') - \chi = 0, \tag{6.20}$$

where we have defined the T'-independent dimensionless parameter

$$\chi \equiv \chi(l) := \ln\frac{2m_e c^2 \xi' l}{I^2} + j, \tag{6.21}$$

which we note has value $\gtrsim 13.84$ for MicroBooNE, assuming [13, 21, 133]:

- mean excitation energy of argon $I = 188.0\,\mathrm{eV}$;
- detector thickness $l \geq 0.3\,\mathrm{cm}$ (the wire separation); and,
- $\xi' \approx 0.0969\,\mathrm{MeV/cm}$ (using Equation 6.18),

in addition to the parameters defined for Equation 6.18. For LArTPCs, the effective detector thickness l (i.e. the 3D Δx traversed per observed 2D hit) is a function of the hit width and the track direction, both of which may be functions of x. We note that $l \geq 0.3\,\mathrm{cm}$ at MicroBooNE and assume that the piece of track considered is short enough for changes in direction (and therefore l) to be negligible.

6.3.2 Low-energy approximation

We demonstrated in Section 6.2 that particles of different masses are most distinguishable near the Bragg peak, which means at lower energies as illustrated in Figure 6.5. It

follows, then, that taking a low-energy limit could be fruitful for particle identification. Taking Taylor expansions about $T' = 0$ under the low-energy approximation $T' \ll 1$,

$$1 - \frac{1}{(T'+1)^2} = 2T' - 3T'^2 + \mathcal{O}\left(T'^3\right) \tag{6.22}$$

and

$$\ln\left(1+T'\right) = T' - \frac{1}{2}T'^2 + \mathcal{O}\left(T'^3\right). \tag{6.23}$$

Substituting these into 6.19 and neglecting terms $\mathcal{O}\left(T'^3\right)$ and above yields

$$T'\left(3T'-2\right)\frac{mc^2}{\xi'}\frac{\mathrm{d}T'}{\mathrm{d}x} - 2T'^2 + \delta(T') - \chi = 0. \tag{6.24}$$

Omitting $\delta(T')$ in the regime $T' <= 1$ leads to a fractional error in $-\mathrm{d}E/\mathrm{d}x$ of $\lesssim 1\%$ for all track-like particles we will consider at MicroBooNE (see Figure 6.9), so we will neglect it. We may then express 6.24 as

$$T'\left(3T'-2\right)\frac{mc^2}{\xi'}\frac{\mathrm{d}T'}{\mathrm{d}x} - 2T'^2 - \chi = 0. \tag{6.25}$$

Rearranging and taking partial fractions yields

$$\left(\frac{3\chi}{\chi+2T'^2} + \frac{4T'}{\chi+2T'^2} - 3\right)\frac{\mathrm{d}T'}{\mathrm{d}x} = -\frac{2\xi'}{mc^2}. \tag{6.26}$$

This may now be integrated directly:

$$x(T') = k - \frac{3mc^2}{2\xi'}\left[\frac{1}{3}\ln\left(\chi+2T'^2\right) + \sqrt{\frac{\chi}{2}}\arctan\left(\sqrt{\frac{2}{\chi}}T'\right) - T'\right], \tag{6.27}$$

where k is the constant of integration. Applying the boundary condition $x(T'=0) = R$, where R is the range of the particle,

$$x(T';R,m) = R - \frac{3mc^2}{2\xi'}\left[\frac{1}{3}\ln\left(1+\frac{2T'^2}{\chi}\right) + \sqrt{\frac{\chi}{2}}\arctan\left(\sqrt{\frac{2}{\chi}}T'\right) - T'\right], \tag{6.28}$$

which is valid in the domain $0 \leq x \leq R$, $T' \ll 1$.We note that this expression also provides a surjective[7] map from the initial scaled modal kinetic energy T'_0 to the particle range R, since $x(T' = T'_0) = 0$ by construction:

$$R = \frac{3mc^2}{2\xi'}\left[\frac{1}{3}\ln\left(1+\frac{2T'^2_0}{\chi}\right) + \sqrt{\frac{\chi}{2}}\arctan\left(\sqrt{\frac{2}{\chi}}T'_0\right) - T'_0\right]. \tag{6.29}$$

[7]The lack of bijectivity in generality (even given $T' \geq 0$) stems from the low-energy approximation; in its domain of applicability, we expect the map to be bijective.

We cannot invert Equation 6.28 analytically but it is possible to construct the function $T' \equiv T'(x)$ efficiently using Newton's method, obviating the need for numerical integration:

$$0 = f(T'; R, m, x) :=$$
$$x - R + \frac{3mc^2}{2\xi'}\left[\frac{1}{3}\ln\left(1 + \frac{2T'^2}{\chi}\right) + \sqrt{\frac{\chi}{2}}\arctan\left(\sqrt{\frac{2}{\chi}}T'\right) - T'\right], \quad (6.30)$$

$$\therefore \frac{\partial f}{\partial T'} = \frac{mc^2}{2\xi'}\left(\frac{3\chi + 4T'}{\chi + 2T'^2} - 3\right). \quad (6.31)$$

Care must be taken to choose an initial T' estimate that robustly converges to the positive root. Given T', we can then calculate $-\mathrm{d}E/\mathrm{d}x$ using Equation 6.25 and the fact that $-\mathrm{d}E/\mathrm{d}x = -\mathrm{d}T/\mathrm{d}x = -mc^2 \cdot \mathrm{d}T'/\mathrm{d}x$:

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \xi'\frac{2T'^2 + \chi}{T'(2 - 3T')}. \quad (6.32)$$

However, greater accuracy (including avoidance of the unphysical pole at $T' = 2/3$; see Figure 6.10) can be achieved by using the exact form for $-\mathrm{d}E/\mathrm{d}x$ from Equation 6.14 in the absence of density effect corrections. Noting $\gamma = T' + 1$ and $\beta^2 = 1 - 1/(T' + 1)^2$,

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \xi'\left(\frac{(T' + 1)^2}{T'(T' + 2)}\left[2\ln\left(1 + T'\right) + \chi\right] - 1\right). \quad (6.33)$$

A comparison of these two expressions is shown in Figure 6.10. Note that Equation 6.31 is still used in both cases to calculate T'; the exact form in Equation 6.33 is only applied in the final step to map $T' \to -\mathrm{d}E/\mathrm{d}x$.

By taking a less precise $T' \ll 1$ expansion during the above derivation, it is also possible to derive an alternative, explicit expression for $-\mathrm{d}E/\mathrm{d}x$ in terms of x:

$$-\frac{\mathrm{d}E}{\mathrm{d}x} \simeq \frac{1}{2}\sqrt{\frac{\xi'\chi mc^2}{R - x}}. \quad (6.34)$$

A derivation of this is given in subappendix B.

Any method designed to exploit these models will likely apply it over some predetermined range R from the end of the particle. R must be chosen such that we are in the domain of our approximation for all track-like particles we expect to encounter. The relationship between the scaled modal kinetic energy T' and the residual range is shown in Figure 6.11 at different masses. The true kinetic energy is not a function of l but

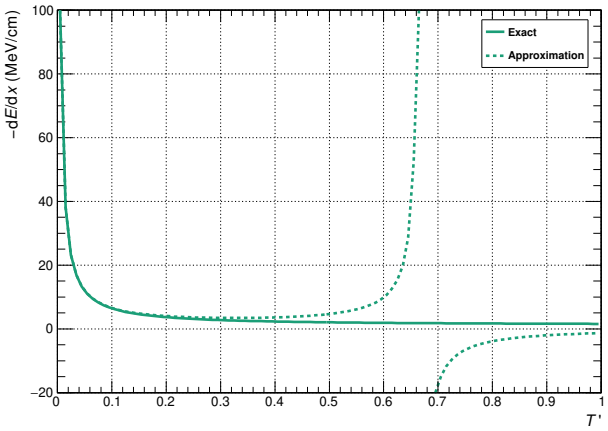


Figure 6.10: Comparison of the 'exact' modal $-\mathrm{d}E/\mathrm{d}x$ in the absence of density effect corrections (Equation 6.33) and its implied second-order $T' \ll 1$ approximation (Equation 6.32). We will use the exact expression for greater accuracy.

this modal proxy is, as described in Section 6.2.3. The domain of applicability of our approximation is similarly a function of l. The approximate modal residual ranges at a selection of T' values are given in Table 6.1 for a variety of particles. For this study, we choose $0.3\,\mathrm{cm} \leq l < 1\,\mathrm{cm}$ (which is satisfied by the vast majority of hits observed) and $R = 10\,\mathrm{cm}$ to balance minimizing T' with maximizing the amount of data to fit.

6.3.3 Validating the $\mathrm{d}E/\mathrm{d}x$ models

These models are validated on MicroBooNE simulation in Chapter 7, using a model for recovering $-\mathrm{d}E/\mathrm{d}x$ from observed charge depositions. Here we use the exact form for the mode given in Equation 6.4 to estimate the error accrued by neglecting density corrections and taking the expansion about $T' = 0$. The results for the explicit expression for $-\mathrm{d}E/\mathrm{d}x$ given by Equation 6.34 (the first-order approximation in T') and the implicit expression given by Equations 6.28 and 6.33 (the second-order approximation in T') are shown in Figure 6.12 for effective detector thickness $l = 0.3\,\mathrm{cm}$, and Figure 6.13 for effective detector thickness $l = 1\,\mathrm{cm}$. As expected, the second-order approximation provides a closer approximation to the true value of the mode. Whilst both methods are fast, whether the extra computational burden of using Newton's method is warranted depends to what degree this improvement in accuracy maps to a concomitant

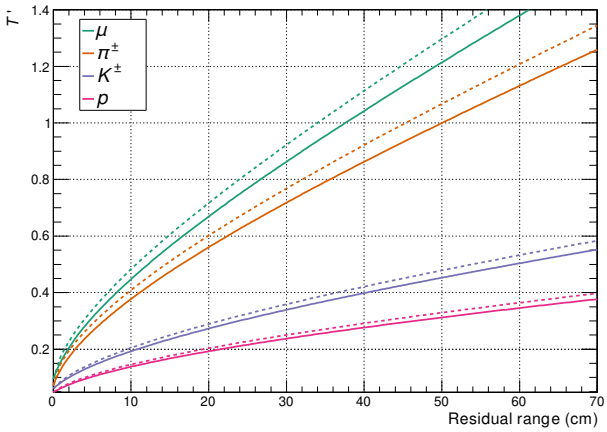


Figure 6.11: The scaled modal kinetic energy T' as a function of residual range for simulated modal particles at MicroBooNE, with $l = 0.3\,\text{cm}$ (solid) and $l = 1\,\text{cm}$ (dashed). The Bethe applicability requirement $0.1 \lesssim \beta\gamma \lesssim 1000$ means we cannot model particles below a certain T', leading to a small offset compared with the true residual range.

Table 6.1: Estimated modal residual ranges in cm at various values of scaled modal kinetic energy T' and effective detector thickness l under MicroBooNE conditions. Note that the values are discretized by l and the true kinetic energy is not a function of l.

		Scaled modal kinetic energy T'				
l (cm)		0.1	0.25	0.5	0.75	1
0.03	μ	0.3	3.3	12.3	24.3	37.8
	$\pi^{\pm}$	0.6	4.8	16.5	32.4	50.1
	$K^{\pm}$	2.1	17.1	59.4	114.9	178.2
	p	4.8	33.3	114.0	219.6	339.6
0.3	μ	1.0	3.0	11.0	22.0	34.0
	$\pi^{\pm}$	1.0	4.0	15.0	29.0	46.0
	$K^{\pm}$	2.0	16.0	54.0	106.0	164.0
	p	5.0	30.0	105.0	202.0	313.0

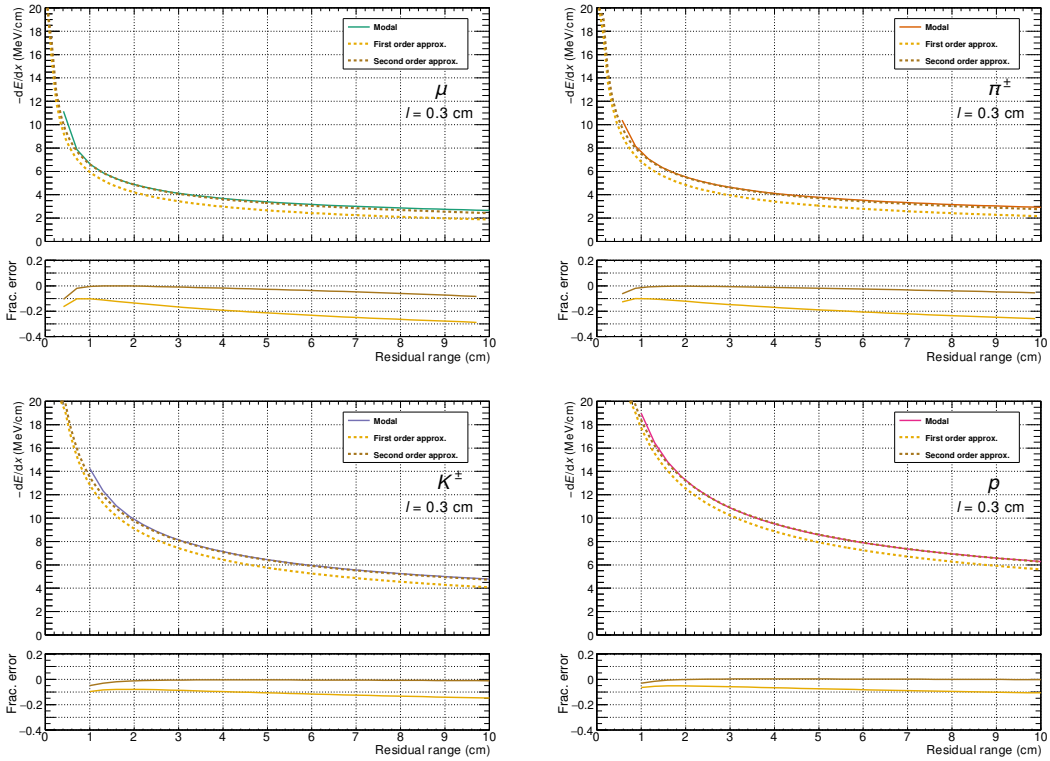


Figure 6.12: A comparison between the modal $-\mathrm{d}E/\mathrm{d}x$ and the first- and second-order approximations described in the main text. Shown are simulated muons (top left), charged pions (top right), charged kaons (bottom left) and proton (bottom right) in a LArTPC with $l = 0.3\,\mathrm{cm}$.

improvement in mass estimation performance. This will be explored in Section 7.6.

6.4 Modelling $\mathrm{d}Q/\mathrm{d}x$

For a particle losing energy solely through ionization, the magnitude of the charge deposition over some Δx is proportional to the amount of energy deposited by the particle over that distance:

$$\Delta Q = \frac{C}{W_{\mathrm{ion}}} \Delta E, \tag{6.35}$$

where

- ΔQ is the deposited charge in units ADC;
- ΔE is the deposited energy in MeV;

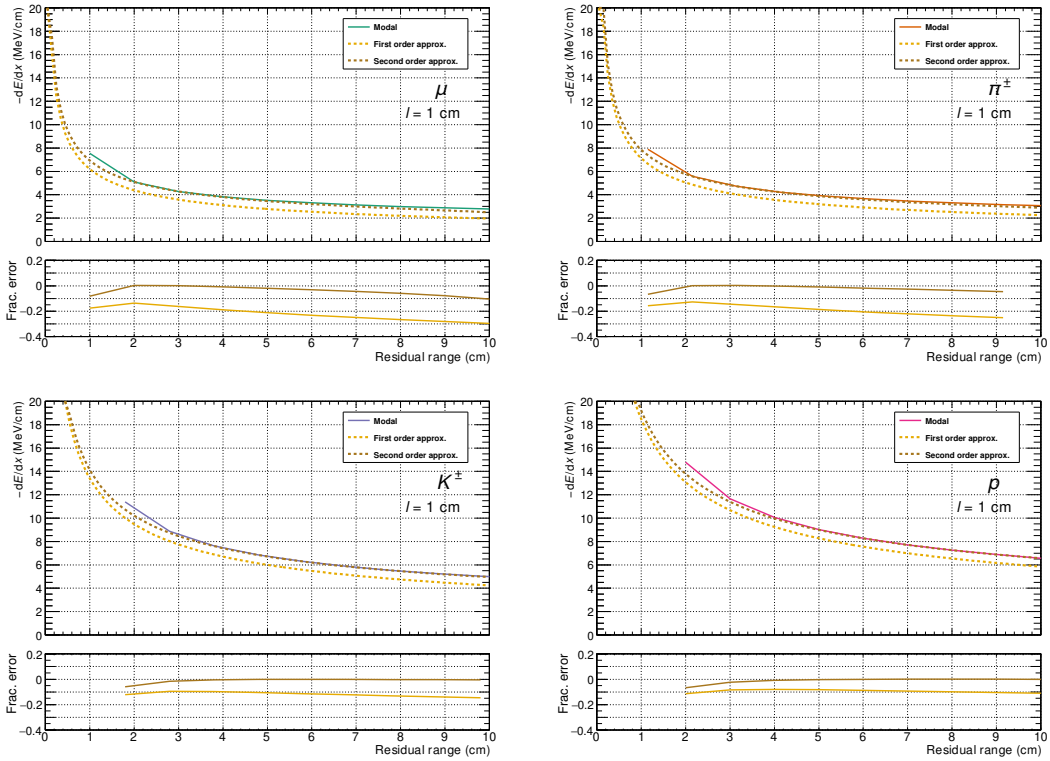


Figure 6.13: A comparison between the modal $-\mathrm{d}E/\mathrm{d}x$ and the first- and second-order approximations described in the main text. Shown are simulated muons (top left), charged pions (top right), charged kaons (bottom left) and proton (bottom right) in a LArTPC with $l = 1\,\mathrm{cm}$.

- W_{ion} is the ionization energy of liquid argon, which is approximately 23.6×10^{-6} MeV/electron; and,
- C is the charge calibration constant at MicroBooNE, which has value of approximately $(5.076 \pm 0.001) \times 10^{-3}$ ADC/e.

In the continuous limit, we may then write[8]

$$\frac{\mathrm{d}Q}{\mathrm{d}x} = \frac{C}{W_{\mathrm{ion}}} \frac{\mathrm{d}E}{\mathrm{d}x}. \tag{6.36}$$

However, the magnitude ΔQ and location x of the *observed* charge is subject to a variety of physical effects that characterise the calorimetric signature observed in the detector

[8]For the sake of consistency with the references, $\mathrm{d}E/\mathrm{d}x$ in this section refers to the rate of energy deposition in the detector, rather than the rate of change of particle energy; the two are related by a minus sign. Note that therefore $\mathrm{d}E/\mathrm{d}x \geq 0$.

for an incident particle. Below is a list of these effects, detailed studies of which can be found in [136] and [137].

- **Electron-ion recombination:** Ionization electrons recombining with their associated argon ions leads to a suppression of observed charge. The magnitude of this effect is sensitive to the amount of charge deposited, the size of the drift electric field, and the density of the detector medium.

- **Space charge effect:** A significant number of cosmic rays (of the order of 10) are observed in the detector during every event readout window. The accumulation of argon ions leads to position-dependent drift electric field distortion which, in turn, affects the magnitude of the recombination effect as described above. This leads to ionization electrons being pulled towards the middle of the detector and a resultant offset in observed hit positions. This effect has been well studied by MicroBooNE; see [138] for a detailed account.

- **Diffusion:** Diffusive processes lead the cloud of ionization electrons to spread out longitudinally and transversely to the drift direction. Longitudinal diffusion leads to a wider signal for longer drift distances and transverse diffusion can cause 'leaking' of charge to neighbouring wires.

- **Electron attenuation:** Electrons drifting for a longer time are more likely to be captured by electronegative contaminants in the argon (such as H_2O and O_2). This means that hits nearer the cathode (i.e. larger drift-time-coordinate) will induce a relatively smaller amount of charge on the collection wire. Due to the purity of argon at MicroBooNE, this effect is relatively small [139].

- **Hardware issues:** ASIC misconfiguration and cross-connected/touching wires distort the electric field between wire planes, leading to a position-dependent response to charge deposition [118].

- **Temporal changes:** Running conditions, such as temperature and argon purity, change over time. These changes affect the detector response to incident charge, the most significant effect being argon purity variation.

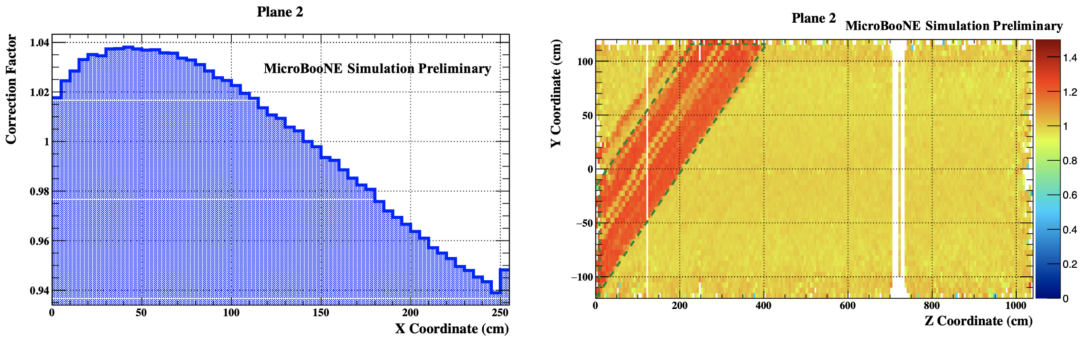


Figure 6.14: The charge calibration values as a function of the drift-time (x) coordinate (left) and of the y-z coordinates. The white area in the right plots shows regions of the detector with unresponsive wires, whereas the redder regions shows areas with misconfigured or cross-connected wires. Figure from [136].

- **Unresponsive wires:** Unresponsive wires in MicroBooNE make parts of the 2D planes insensitive to ionization electrons, which can lead to lost hits if information is missing or ambiguous in corresponding positions in the other 2D planes. This is taken into account in the definition of MicroBooNE fiducial region, which aims to avoid the worst-affected areas.

These corrections form a map that allows us to infer $\mathrm{d}E/\mathrm{d}x$ from observed $\mathrm{d}Q/\mathrm{d}x$. The space charge effect correction is described in Chapter 3. We will consider the recombination correction in particular detail, and term the remaining effects 'charge calibration'. Extensive calorimetric calibration work has been carried out at MicroBooNE [136, 137] to understand and correct for these effects. The resulting charge correction factors as a function of position in the detector are shown in Figure 6.14. Following [109], we use an analytic approximation to the drift-time coordinate correction factor and, for the y-z plane correction factor, a 2D step function in which the region demarcated by green dash lines is corrected by a factor equal to the mean factor in this region, and 1 elsewhere (i.e. no correction).

6.4.1 Modelling recombination

Charge depositions associated with track-like particles in MicroBooNE experience significant suppression due to electron-ion recombination. The size of the suppression is a function of the local charge density, as well as detector parameters. To keep the treat-

ment clear, we will compare the *observed* charge deposition rate $\mathrm{d}Q/\mathrm{d}x$ with the *true* energy deposition rate $\mathrm{d}E/\mathrm{d}x$. In the absence of recombination, Equation 6.36 then holds true.

To make this more concrete, we define the charge suppression factor

$$R := \frac{W_{\text{ion}}}{C}\frac{\mathrm{d}Q/\mathrm{d}x}{\mathrm{d}E/\mathrm{d}x} \tag{6.37}$$

such that $R = 1$ corresponds to the no-recombination case of equation 6.36. In the models we consider, R can be expressed as a function of either $\mathrm{d}Q/\mathrm{d}x$ or $\mathrm{d}E/\mathrm{d}x$. The factor $\frac{C}{W_{\text{ion}}}\frac{\mathrm{d}E}{\mathrm{d}x}$ can be seen to correspond to the 'expected' charge deposition in the absence of recombination. The map from stopping power to observed charge depositions

$$\frac{\mathrm{d}Q}{\mathrm{d}x} = R\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)\frac{C}{W_{\text{ion}}}\frac{\mathrm{d}E}{\mathrm{d}x} \tag{6.38}$$

is therefore of interest when generating particles; during reconstruction, we are more interested in the inverse map (from observed charge depositions to stopping power)

$$\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{1}{R\left(\mathrm{d}Q/\mathrm{d}x\right)}\frac{W_{\text{ion}}}{C}\frac{\mathrm{d}Q}{\mathrm{d}x}, \tag{6.39}$$

where we are now expressing R as as function of the observed charge deposition rate $\mathrm{d}Q/\mathrm{d}x$.

To model this effect, there are two canonical options:

1. **Birks' law:** An empirical formula modelling the recombination effect and developed for organic scintillators:

$$R_{\text{Birks}}\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right) = \frac{A}{1 + \frac{k}{\rho\mathcal{E}}\frac{\mathrm{d}E}{\mathrm{d}x}}, \tag{6.40}$$

   where

   - A is a fit parameter, measured at ICARUS to be 0.800 ± 0.003;
   - k is another fit parameter, measured at ICARUS to be $(0.0468 \pm 0.0006)\ \frac{\mathrm{g/cm^2}}{\mathrm{MeV}} \cdot \frac{\mathrm{kV}}{\mathrm{cm}}$;
   - ρ is the density, as defined for Equation 6.2 (and taking value $\sim 1.4\,\mathrm{g\,cm^{-3}}$ for MicroBooNE); and,

- $\mathcal{E}$ is the electric field strength in the MicroBooNE detector, which is approximately 0.273 keV/cm.

The map from $\mathrm{d}Q/\mathrm{d}x$ to $\mathrm{d}E/\mathrm{d}x$ can then be calculated using Equation 6.39:

$$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{Birks}} = \frac{\frac{W_{\text{ion}}}{C}\frac{\mathrm{d}Q}{\mathrm{d}x}}{A - \frac{k}{\rho\mathcal{E}} \cdot \frac{W_{\text{ion}}}{C}\frac{\mathrm{d}Q}{\mathrm{d}x}}, \tag{6.41}$$

which exhibits an unphysical pole at $\mathrm{d}Q/\mathrm{d}x = A\rho\mathcal{E}C/kW_{\text{ion}} \approx 1405\,\text{ADC/cm}$ for MicroBooNE.

2. **Modified box model[9]:** A model of recombination developed at ArgoNeuT (a smaller LArTPC) [140] that is tuned to resemble Birks' law in a useful $\mathrm{d}E/\mathrm{d}x$ domain but does not exhibit a pole in the resulting inverse map:

$$R_{\text{ModBox}}\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right) = \frac{\ln\left(A + \frac{B}{\rho\mathcal{E}}\frac{\mathrm{d}E}{\mathrm{d}x}\right)}{\frac{B}{\rho\mathcal{E}}\frac{\mathrm{d}E}{\mathrm{d}x}}, \tag{6.42}$$

where:

- $A = 0.930$ and $B = 0.212\ \frac{\text{g/cm}^2}{\text{MeV}} \cdot \frac{\text{kV}}{\text{cm}}$ are fit parameters whose fit parameters measured at ArgoNeuT under an electric field of 0.481 kV/cm [140], and
- ρ and $\mathcal{E}$ are as defined for Equation 6.40.

We note that the Birks' law and modified box model parameters have recently been refit for MicroBooNE [141]. Using Equation 6.39, the map from $\mathrm{d}Q/\mathrm{d}x$ to $\mathrm{d}E/\mathrm{d}x$ is then

$$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{ModBox}} = \frac{\rho\mathcal{E}}{B}\left[\exp\left(\frac{B}{\rho\mathcal{E}} \cdot \frac{W_{\text{ion}}}{C}\frac{\mathrm{d}Q}{\mathrm{d}x}\right) - A\right]. \tag{6.43}$$

The models are compared in Figure 6.15. Both models exhibit physically-aberrant behaviour at large $\mathrm{d}E/\mathrm{d}x$ values. Birks' law implies a constant amount of charge deposition in the limit of large $\mathrm{d}E/\mathrm{d}x$ (corresponding to a pole in the inverse map), whereas the modified box model exhibits a unphysical suppression of charge deposition at large $\mathrm{d}E/\mathrm{d}x$ (leading to exponential increase in the inverse map). The difference between the

[9] An alternative version of the modified box model where $\mathcal{E} \to \mathcal{E}\sin\theta$ to account for angular dependence is also used but is not the default implementation at MicroBooNE. θ here is the angle between the electric field and the particle's direction of motion.

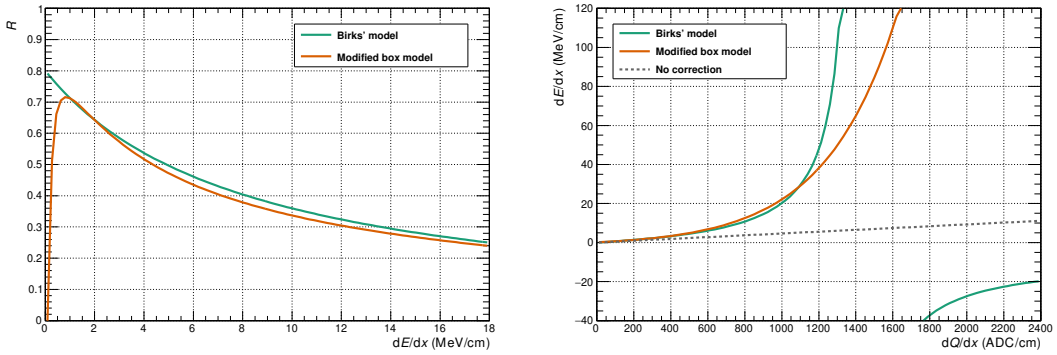


Figure 6.15: Comparison of the suppression factors (left) and the maps from observed $\mathrm{d}Q/\mathrm{d}x$ to true $\mathrm{d}E/\mathrm{d}x$ (right) for Birks' model and the modified box model as a function of the true deposited energy $\mathrm{d}E/\mathrm{d}x$ in a MicroBooNE-like detector (i.e. a LArTPC with MicroBooNE fits for Birks' law/ModBox parameters). Note the pole in Birks' model map at $\mathrm{d}Q/\mathrm{d}x \approx 1405\,\mathrm{ADC/cm}$.

two models in the region $40\,\mathrm{MeV/cm} \lesssim \mathrm{d}E/\mathrm{d}x \lesssim 100\,\mathrm{MeV/cm}$ is significant, and this is a region in which we have some interest (see e.g. Figure 6.10). The pole in Birks' model is manifestly unphysical, as is the exponential increase suggested by the modified box model at large energies. Another study suggests a correction to Birks' law leading to linear behaviour at large $\mathrm{d}E/\mathrm{d}x$ values [142].

The upper limit of $\mathrm{d}E/\mathrm{d}x$ at which each of these two models become unphysical is as yet unclear. The liquid argon community seems to prefer the modified box model so, in the absence of further information, we will use it too. However, in light of the uncertainty of the models' relative accuracies, we will require $150\,\mathrm{ADC/cm} \leq \mathrm{d}Q/\mathrm{d}x \leq 1200\,\mathrm{ADC/cm}$ such that the two models only deviate minimally. Any hits with $\mathrm{d}Q/\mathrm{d}x$ values outside this range will be considered outliers and not be used to fit $\mathrm{d}E/\mathrm{d}x$.

Combining the modified box model (Equation 6.42) with our first-order model for $\mathrm{d}E/\mathrm{d}x$ (Equation 6.32, noting the $\mathrm{d}E/\mathrm{d}x$ sign change), yields the first-order model for $\mathrm{d}Q/\mathrm{d}x$:

$$\frac{\mathrm{d}Q}{\mathrm{d}x} = \frac{\rho\mathcal{E}C}{BW_{\mathrm{ion}}} \ln\left(A + \frac{B}{2\rho\mathcal{E}}\sqrt{\frac{\xi'\chi mc^2}{R-x}}\right). \tag{6.44}$$

Similarly, combining the modified box model with the second-order $\mathrm{d}E/\mathrm{d}x$ model (Equa-

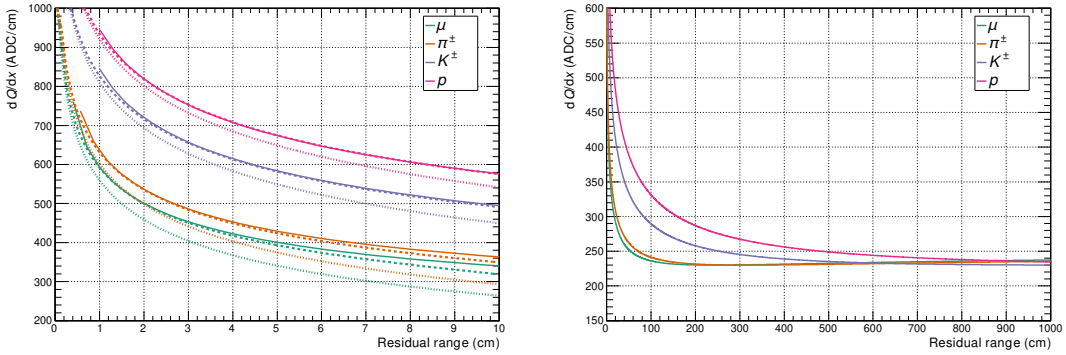


Figure 6.16: Left: comparison of the exact (solid, Equation 6.4), second-order (dashed, Equations 6.45 and 6.28), first-order (dotted, Equation 6.44) modal $\mathrm{d}Q/\mathrm{d}x$ expressions using the modified box model. The 10 cm range is as discussed in Section 6.3.2. Right: the exact modal $\mathrm{d}Q/\mathrm{d}x$ for muons, charged pions, charged kaons and protons using the modified box model over a larger residual range, as per Figure 6.5.

tion 6.33) yields

$$\frac{\mathrm{d}Q}{\mathrm{d}x} = \frac{\rho\mathcal{E}C}{BW_{\mathrm{ion}}} \ln\left[A + \frac{B\xi'}{\rho\mathcal{E}}\left(\frac{(T'+1)^2}{T'(T'+2)}\left[2\ln\left(1+T'\right)+\chi\right]-1\right)\right], \tag{6.45}$$

where the scaled modal kinetic energy values $T' \equiv T'(x)$ are calculated via the method proposed in Section 6.3.2.

6.4.2 Application to particle identification

The existence of a low-error closed-form expression for $\mathrm{d}Q/\mathrm{d}x$ (Equation 6.45) with an implicit dependence on particle mass has direct application to particle identification. Working with the mode, rather than mean, of the energy loss rate distributions was the key insight that enabled the derivation, but this requires being able to robustly estimate the mode from the real data. This will be explored in Chapter 7, where we develop a robust new method for performing particle identification using the derivations from this chapter.

Chapter 7

Particle identification using modal $\mathrm{d}Q/\mathrm{d}x$

This chapter introduces a new theoretically motivated particle identification method (FOMA) based on a first-order modal approximation to low-energy $\mathrm{d}Q/\mathrm{d}x$ developed in the previous chapter. The new method and an algorithm to reconstruct three-dimensional $\mathrm{d}E/\mathrm{d}x$ are described and compared with PIDA, a standard PID method used in MicroBooNE. The MIP/proton discrimination ability of FOMA is shown to be superior to that of PIDA despite requiring no parameter tuning. The robustness of both FOMA and PIDA to MC/data disagreement is explored.

7.1 Introduction

High-quality particle identification (PID) is necessary for accurate energy reconstruction, which in turn feeds directly into MicroBooNE's primary goals. This is usually separated into two distinct tasks: track/shower classification and track PID. The former has a strong bearing on the reconstruction, as showers produce more disperse distributions of hits that require different clustering methods to the well-defined trails produced by tracks. Track/shower classification is therefore a feature internal to Pandora, where it is performed using machine learning techniques with a set of physics-inspired features. In this study, 'showers' refers to electromagnetic showers produced by electrons and

photons, and 'tracks' to all other charged particles. At MicroBooNE, tracks primarily comprise muons, charged pions, protons and (less frequently) kaons. Additionally, low-energy electrons can appear track-like if they do not have enough energy to produce an electromagnetic shower, and are often misclassified as such.

PID does not currently feature inside Pandora as it lies outside the remit of the pattern recognition stage. Only a few examples of Pandora reconstruction for LArTPCs use calorimetric information, one of them being the vertex selection algorithm (see Chapter 5) and another the 3D track/shower identification algorithm. Beyond pattern recognition, both track/shower classification and track PID are critical to reconstructing particle (including neutrino) energies; shower energies are commonly reconstructed by calibrating their total charge, whereas track energies require hit-level correction for recombination effects and are preferably reconstructed from range rather than on a per-hit basis. This necessitates accurate knowledge of the particle mass and therefore the particle ID.

Improving track PID at MicroBooNE, which we will refer to as simply PID, is the subject of the next two chapters. The most common variable used to assist PID is PIDA, a scalar variable derived from a phenomenological model tuned on MC simulation. In this chapter, we present a more theoretically rigorous but equally flexible approach called FOMA and demonstrate that it generally performs better than PIDA in distinguishing MIPs and protons. Both models require good three-dimensional $\mathrm{d}E/\mathrm{d}x$ reconstruction to be effective, so we begin by developing an algorithm for that task.

7.2 Measuring $\mathrm{d}Q/\mathrm{d}x$

Accurately measuring $\mathrm{d}Q/\mathrm{d}x$ and ascribing a 3D position to each measurement is non-trivial. We consider for each 2D hit i an inferred 3D position $\boldsymbol{x}_i$, a 3D path length Δx_i and a charge deposition ΔQ_i, such that the rate of charge deposition at hit i is approximately $\Delta Q_i/\Delta x_i$. These tuples of values for every 2D hit along a track in a given plane together form a data set $\mathcal{X} := \{\boldsymbol{x}_i, \Delta Q_i, \Delta x_i\}_{i\in[1,N]}$ which can be used to reconstruct 3D $\mathrm{d}Q/\mathrm{d}x$ or $\mathrm{d}E/\mathrm{d}x$ using the method described in Section 6.4.

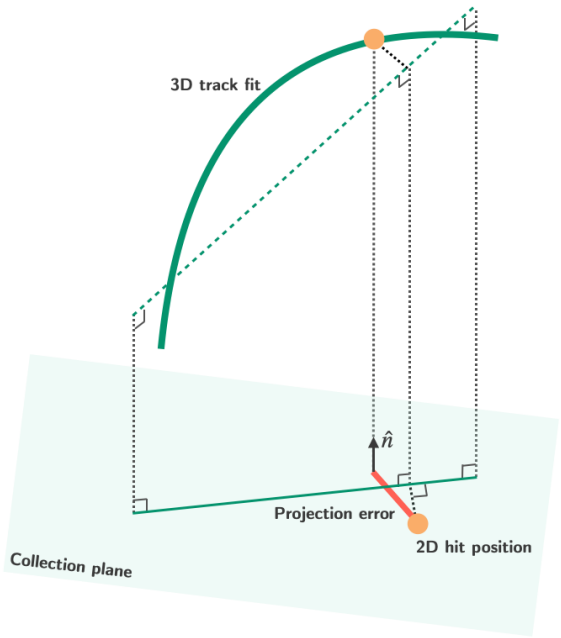


Figure 7.1: The geometry of the 2D-to-3D hit projection. A 3D track fit axis (green dotted line) is projected into the collection plane. The 2D hit (orange circle) in the collection plane is first projected onto the 2D projected track axis, then onto the 3D track fit axis, then onto the 3D track fit. The inferred 3D hit position can then be projected back into the collection plane and its error measured. $\hat{n}$ in the unit normal to the collection plane.

7.2.1 Charge deposition

The charge deposition value we will use for Q_i is integrated ADC count, a measure of the response of the detector to deposited charge. This is provided on a per-hit basis by the hit-finding algorithm for each 2D hit. We choose to use the collection plane (W) hits, which provide the best calorimetric resolution (see Chapter 4); improved signal processing methods in the future will enable incorporation of induction plane charge.

7.2.2 3D hit position

Calculating the 3D hit position $\boldsymbol{x}_i$ and the 3D path length Δx_i is less straightforward. As already mentioned, these are 2D collection plane hits, so we must infer their 3D position using properties of the full 3D reconstruction.

We may calculate $\boldsymbol{x}_i$ in two distinct ways and compare the results:

1. Most collection plane hits will have an associated 3D hit, whose position is obtained via a process of matching 2D hits from different views and 3D trajectory-smoothing. We then map from the 2D to the 3D hit (if one exists) to obtain its 3D position vector. 2D hits that cannot be matched in other views will not be reconstructed into a 3D hit (see Chapter 4).

2. Even if a 2D hit has no corresponding 3D hit, we may perform a sliding linear fit to the associated PFO's 3D hits, project the fit axis into the collection plane, find the closest approach between the 2D hit and the projected fit, and use the equivalent point on the 3D fit to obtain an approximate position. This is demonstrated in Figure 7.1.

A convenient aspect of this process is that the inferred 3D hit position may be projected back into the collection plane, allowing for an positional error to be calculated for both simulation and data. Further, hits whose projection error is large (say $>$ 5cm in 2D-projected distance) can be considered spurious and excluded from certain kinds of analysis, including the PID methods developed in this thesis. The cause of such outliers can include, for instance, a δ-ray hit mis-clustered into a track. This calculation is subject to the space charge effect correction described in Section 6.4.

7.2.3 3D path length

Finally, to calculate Δx_i we need to use properties of both the 2D hit and its associated 3D fit. We will first calculate for hit i the 2D projected distance in the wire plane, $L_{xz,i}$ then extend this to the true 3D path length, Δx_i. We begin with two assumptions which are empirically well-satisfied:

1. Collection plane hits are mostly contiguous in their wire plane if they represent a continuous piece of well-reconstructed track, i.e. they do not consistently overlap or have gaps between them. Detector gaps or otherwise missing hits are not an issue here as we are not able to reconstruct the lost hits in any case.

2. The widths of adjacent hits are similar. This is mostly a function of the track angle at the position of a given hit and satisfied by the largely linear nature of tracks at MicroBooNE.

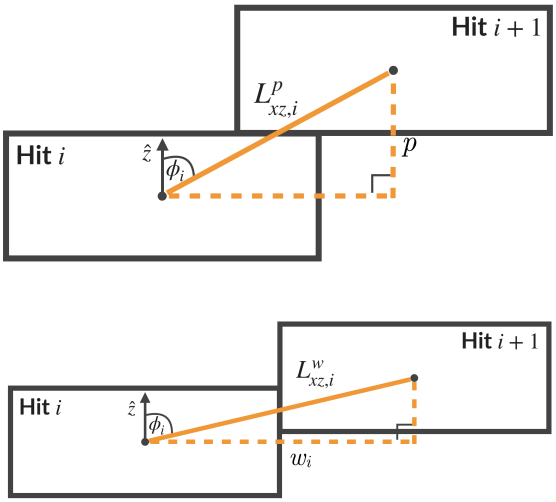


Figure 7.2: Top: the geometry of two top-to-bottom contiguous hits in the collection plane. Their inter-centroid distance $L^p_{xz,i}$ is a function of the angle ϕ_i and the wire pitch p. Bottom: the geometry of two side-to-side contiguous hits in the collection plane. Their inter-centroid distance $L^w_{xz,i}$ is a function of the angle ϕ_i and the hit width w_i.

Each 2D hit i has a variable width w_i, determined by the width of the signal processed by the hit-finder, and a fixed height p, corresponding to the wire pitch. The width therefore extends in the drift time coordinate axis (the x-direction) and, in the collection plane, the pitch extends in the z-direction. We consider two adjacent hits, i and $i+1$, in the collection plane and define an angle ϕ_i to be the acute angle between the vector between the hits' centroids and the z-axis.

Due to the rectangular shape of the hits, there are two geometrically distinct cases of two hits just touching:

1. **Top-to-bottom:** As shown at the top of Figure 7.2. From trigonometry:

$$L^p_{xz,i} = \frac{p}{\cos\phi_i} \tag{7.1}$$

2. **Side-to-side:** As shown at the bottom of Figure 7.2. From trigonometry:

$$L^w_{xz,i} = \frac{w_i}{\sin\phi_i} \tag{7.2}$$

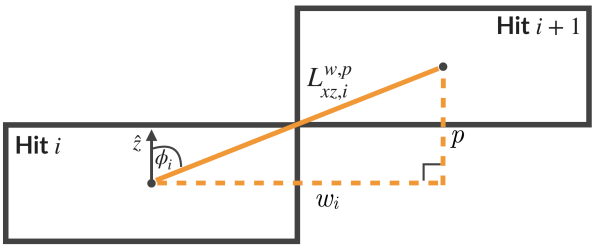


Figure 7.3: The geometry of two corner-to-corner contiguous hits in the collection plane. This forms a boundary case between the geometries in Figure 7.2.

There is a boundary condition between these solutions: they must yield the same value of $L_{xz,i}$ when the hits are corner-to-corner (Figure 7.3). This occurs when

$$\tan\phi_i = \frac{w_i}{p}. \tag{7.3}$$

Then, from 7.1:

$$\begin{aligned} L^p_{xz,i}\Big|_{\tan\phi_i=\frac{w_i}{p}} &= \frac{p}{\cos\left[\arctan\left(\frac{w_i}{p}\right)\right]} \\ &= \sqrt{w_i^2+p^2}, \end{aligned} \tag{7.4}$$

where we have made use of the result $\cos(\arctan x) = \frac{1}{\sqrt{x^2+1}}$. This value can be confirmed using Pythagoras' theorem. Similarly, from 7.2 and using the result $\sin(\arctan x) = \frac{x}{\sqrt{x^2+1}}$,

$$\begin{aligned} L^w_{xz,i}\Big|_{\tan\phi=\frac{w_i}{p}} &= \frac{w_i}{\sin\left[\arctan\left(\frac{w_i}{p}\right)\right]} \\ &= \sqrt{w_i^2+p^2} \end{aligned} \tag{7.5}$$

$$\therefore L^p_{xz,i}\Big|_{\tan\phi_i=\frac{w_i}{p}} = L^w_{xz,i}\Big|_{\tan\phi_i=\frac{w_i}{p}} \tag{7.6}$$

as required.

We can derive a rule for choosing between the two schema by considering d_p, i, the component of the inter-hit-centroid distance parallel to the z-axis for hit i, and d_w, i, the its component parallel to the x-axis for hit i (see Figure 7.4). The simplest rule for choosing the correct measurement is

$$L_{xz,i} = \begin{cases} L^w_{xz,i}, & \text{if } 0 \leq d_p, i \leq p \\ L^p_{xz,i}, & \text{if } 0 \leq d_w, i \leq w_i. \end{cases} \tag{7.7}$$

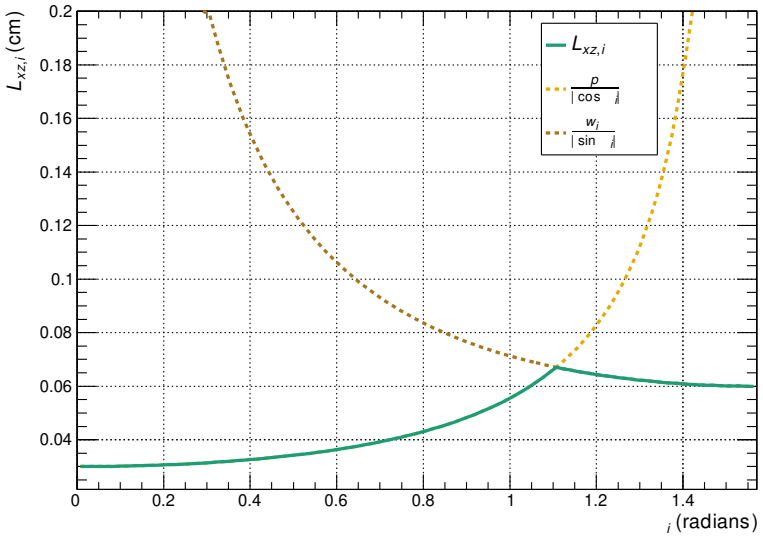


Figure 7.4: The behaviour of the 2D path length function defined in Equation 7.10 as a function of the azimuthal angle ϕ_i, along with its two component functions. For illustrative purposes, we have chosen $p = 0.03\,\text{cm}$ and $w_i = 0.06\,\text{cm}$; the behaviour therefore changes at $\phi_i = \arctan(w_i/p) \approx 1.11\,\text{rad}$.

Using $\tan\phi_i = \frac{d_{w,i}}{d_{p,i}}$ and substituting the definitions of $L^w_{xz,i}$ and $L^p_{xz,i}$ leads to the more useful rule:

$$L_{xz,i} = \begin{cases} \frac{p}{\cos\phi_i}, & \text{if } \tan\phi_i < \frac{w_i}{p} \\ \frac{w_i}{\sin\phi_i}, & \text{otherwise.} \end{cases} \tag{7.8}$$

This gives the projected 2D distance of the piece of track in the azimuthal (x-z) plane. The agreement of the rule with our boundary condition at $\tan\phi_i = \frac{w_i}{p}$ is now manifest. Since we defined ϕ_i such $\phi_i \in \left[0, \frac{\pi}{2}\right]$, the domain of $\tan\phi_i$ is given by

$$\tan\phi_i \in [0, \infty]\,. \tag{7.9}$$

We may relax the definition of ϕ_i for practical purposes such that $\phi_i \in [-\pi, \pi]$ (no longer requiring an acute or positive angle by construction), in which case Equation 7.8 becomes

$$L_{xz,i} = \begin{cases} \frac{p}{|\cos\phi_i|}, & \text{if } |\tan\phi_i| < \frac{w_i}{p} \\ \frac{w_i}{|\sin\phi_i|}, & \text{otherwise.} \end{cases} \tag{7.10}$$

This is represented graphically in Figure 7.4. As the figure makes clear, we can equivalently write this as

$$L_{xz,i} = \min\left\{\frac{p}{|\cos\phi_i|}, \frac{w_i}{|\sin\phi_i|}\right\}. \tag{7.11}$$

We now extend this to the true 3D length, Δx_i. Let the 3D track locally to hit i make an angle θ_i with respect to the polar (y) axis. We add the assumption that the particle travels all the way across each hit, which is a good assumption if the number of hits in the azimuthal (collection) plane $\gg 1$. More concretely, we require

$$\Delta x_i \sin\theta_i \geq \begin{cases} p, & \text{if } |\tan\phi_i| < \frac{w_i}{p} \\ w_i, & \text{otherwise,} \end{cases} \tag{7.12}$$

where $\theta_i \in \left[0, \frac{\pi}{2}\right]$ as per the polar angle of a canonical spherical coordinate system. Then $\Delta x_i \sin\theta_i = L_{xz,i}$, so

$$\Delta x_i = \frac{L_{xz,i}}{\sin\theta_i} \tag{7.13}$$

$$= \begin{cases} \frac{p}{|\cos\phi_i|\sin\theta_i}, & \text{if } |\tan\phi_i| < \frac{w_i}{p} \\ \frac{w_i}{|\sin\phi_i|\sin\theta_i}, & \text{otherwise.} \end{cases} \tag{7.14}$$

Analogously to before, allowing $\theta_i \in \left[-\frac{\pi}{2}, +\frac{\pi}{2}\right]$ and rewriting as a minimum yields the final result:

$$\Delta x_i = \min\left\{\frac{p}{|\cos\phi_i \sin\theta_i|}, \frac{w_i}{|\sin\theta_i \sin\phi_i|}\right\}. \tag{7.15}$$

The poles[1] at $\theta_i = n\pi, n \in \mathbb{Z}$ are avoided using the condition of applicability in Equation 7.12. We will use Equations 7.12 and 7.15 to obtain values of Δx_i for 2D collection plane hits where possible.

7.2.4 Calculating $\mathrm{d}Q/\mathrm{d}x$ and $\mathrm{d}E/\mathrm{d}x$

This calculation is subject to a number of corrections, which are described in detail in 6.4. After correcting for the space charge effect and the charge calibration, we arrive at 3D $\mathrm{d}Q/\mathrm{d}x$. By further applying the recombination correction, we reconstruct 3D $\mathrm{d}E/\mathrm{d}x$. A demonstration of $\{(\mathrm{d}E/\mathrm{d}x)_i, \Delta x_i, w_i\}_{0\leq i<N}$ (which we call the 'calorimetric triplet') is shown in Figure 7.5.

[1] When a particle moves near-parallel to a wire in a given plane, charge is only received in one wire position it is possible only one or a small number of hits are created. This makes 3D reconstruction difficult or impossible.

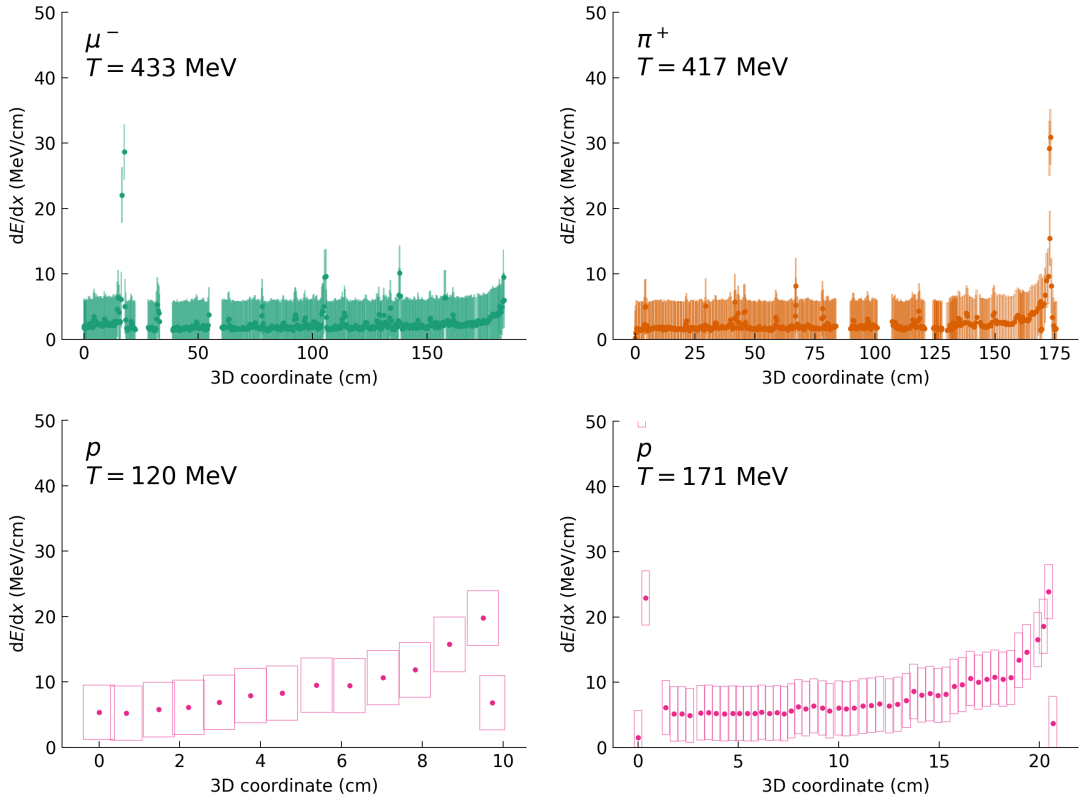


Figure 7.5: A visual representation of the calorimetric triplet $\{(\mathrm{d}E/\mathrm{d}x)_i, \Delta x_i, w_i\}_{0 \leq i < N}$ for four particles, randomly chosen from the a high-purity and -completeness data subset. The cumulative sum of the Δx_i values leads to the 3D coordinate represented on the x-axis. The x-extent of the boxes represents the hit width w_i and the y-extent is arbitrary. Hits with no x-gap nor x-overlap suggest good 3D hit reconstruction, which is largely what is observed here. Detector gaps are visible in some plots. Bragg peaks are clear in all four cases, along with noise due to shared or poorly reconstructed hits.

7.3 PIDA

PIDA is a phenomenological model for characterizing the Bragg peak to perform particle identification [140], widely used at MicroBooNE. The derivation begins by noting that the Bethe-Bloch equation suggests that $\mathrm{d}E/\mathrm{d}x$ is an approximate power law in R, the residual range, at low energies (see Figure 7.6). The creators therefore model $\mathrm{d}E/\mathrm{d}x$ as

$$\frac{\mathrm{d}E}{\mathrm{d}x} = AR^b. \tag{7.16}$$

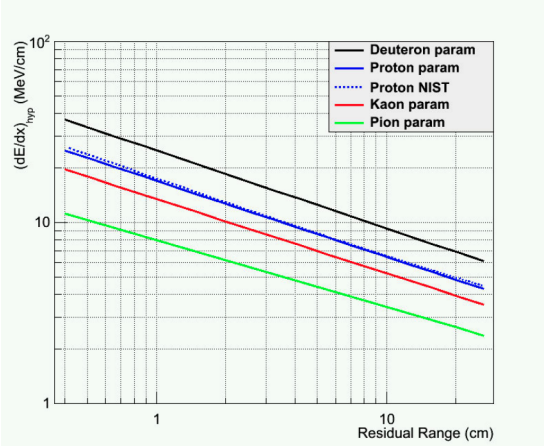


Figure 7.6: A demonstration of the power law dependence of simulated track particles in an ICARUS-like detector at low energies as suggested by the Bethe-Bloch equation, which leads to the definition of the PIDA variable. Figure from [140].

for arbitrary constants A and b in the last $\sim 11\,\text{cm}$ of track. Integrating yields T_{range}, the kinetic energy deposited over the range, noting that $T_{\text{range}}(R=0)=0$:

$$T_{\text{range}} = \frac{A}{b+1} R^{b+1}. \tag{7.17}$$

Performing a fit to protons, pions, muons and kaons reveals that the optimal value of b varies over a small range whereas A varies more substantially. b was therefore set as a number within this range (0.42) and define PIDA as

$$\text{PIDA} := \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{hit}i} R_i^{0.42}. \tag{7.18}$$

Appealingly, particle identification using PIDA is performed graphically by plotting the values in a histogram and looking for peaks directly in the data distribution, as shown in Figure 7.7. From the MC simulation, we know the order of the peaks and the expected position of each particle. Though the $b = 0.42$ decision was tuned using MC simulation, the method now becomes somewhat decoupled from MC/data disagreement. Cuts can be chosen by eye or through a 1D unsupervised clustering method; in practice, PIDA is usually used as just one variable among many in multivariate analyses.

For characterizing the performance of PID methods in this work, we evaluate their performance on primary particles (i.e. the first visible daughters of the neutrino) using simulated BNB events with MC neutrinos generated by GENIE and MC cosmic rays

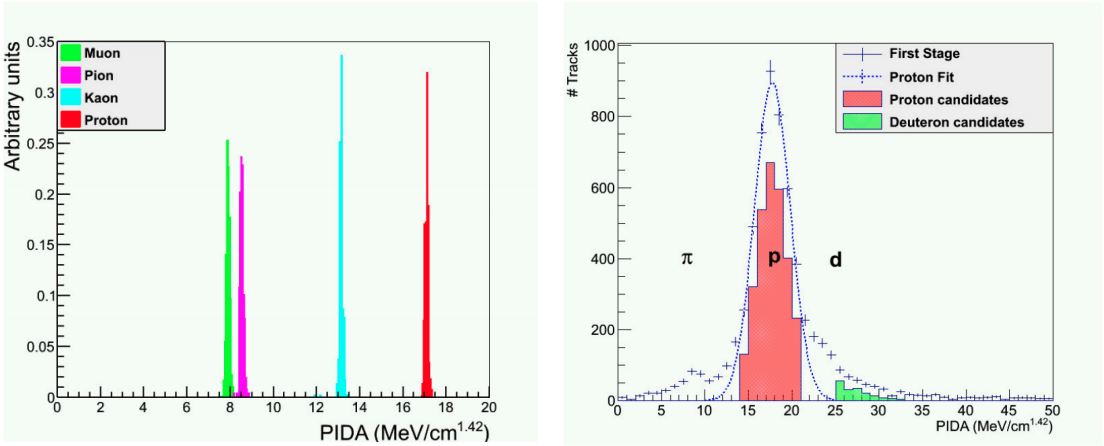


Figure 7.7: Left: an ideal PIDA histogram based on MC truth information and neglecting detector effects, showing almost perfect discriminatory power amongst all particles (including muons and pions). Right: a PIDA histogram based on a real data selection from ICARUS during exposure of 1.35×10^{20} protons-on-target in the NuMI neutrino beam, showing distinguishable peaks but greatly blurred in comparison with the MC truth. Figure from [140].

generated by CORSIKA. See Chapter 3 for details about the procedure for simulating events. We form two subsets of the primary particles in these events: `MC_RECO_TRACKS` and `MC_CLEAN_TRACKS`. The `MC_RECO_TRACKS` data set uses the following criteria:

- the particle was reconstructed by Pandora but not necessarily correctly (see Chapter 4 for a definition of reconstruction correctness);
- the particle has at least 3 collection plane hits whose $\mathrm{d}Q/\mathrm{d}x$ and 3D position can be reconstructed with an error of less than 5 cm compared with the 2D collection plane position (note that this error can be calculated for data too, see Section 7.2.3); and,
- the particle is classified as a track by Pandora.

Together these form a loose set of criteria requiring no MC truth information that place no requirements on the nature of the particle besides it being reconstructed as a track and having enough hits to calculate PIDA/FOMA. Such cuts can therefore be applied to data as well as simulation. The `MC_CLEAN_TRACKS` set additionally makes use of the following criteria:

- the particle has no reconstructed daughters;
- the particle is well-matched to an MC particle (with hit-matching purity and completeness both at least 0.9);
- the particle is a track, based on truth information from the matched MC particle; and,
- at least 90% of the energy-weighted number of particles below this particle in the daughter hierarchy are fully contained within the fiducial volume of the detector, based on truth information from the matched MC particle.

The purpose of these is to create a 'clean' set on which PID should be easier and that can be used for validating PIDA/FOMA, selecting for fully contained, stopping particles with high-quality MC truth information. Before performing the study, we randomly divide each dataset into a training set and a test set, such that no particles from the same event occur in both the test and train set. The same training and test set divisions are used in Chapters 8 and 9 to facilitate unbiased results and a fair comparison of methods. The training set is used to optimize the cut value above which we classify particle as protons, and the test set is only used for validation.

Performance of PIDA on the binary classification task of distinguishing MIPs from protons is shown in Figure 7.8. Optimization of the PIDA MIP/proton cut value is performed using logistic regression, which returns the probabilities used to construct the ROC curve. For the purposes of evaluating PID, we ignore other particles; this condition is relaxed in Chapter 9 in the context of applying these methods in practice. We additionally consider performance on the `MC_RECO_TRACKS` set where the true kinetic energy of the particle is less than 200 MeV and the PID task is more challenging. We refer to this subset of `MC_RECO_TRACKS` as `MC_RECO_TRACKS_LE`. We contrast the performance of classifiers with the baseline classifier, which is defined as the classifier that always chooses the most common particle in the training set. The quality of prediction is highest on the `MC_CLEAN_TRACKS` set (AUC = 0.916) and, as expected, degrades on the `MC_RECO_TRACKS` set (AUC = 0.883) and again on the low-energy `MC_RECO_TRACKS` low-energy data (AUC = 0.849).

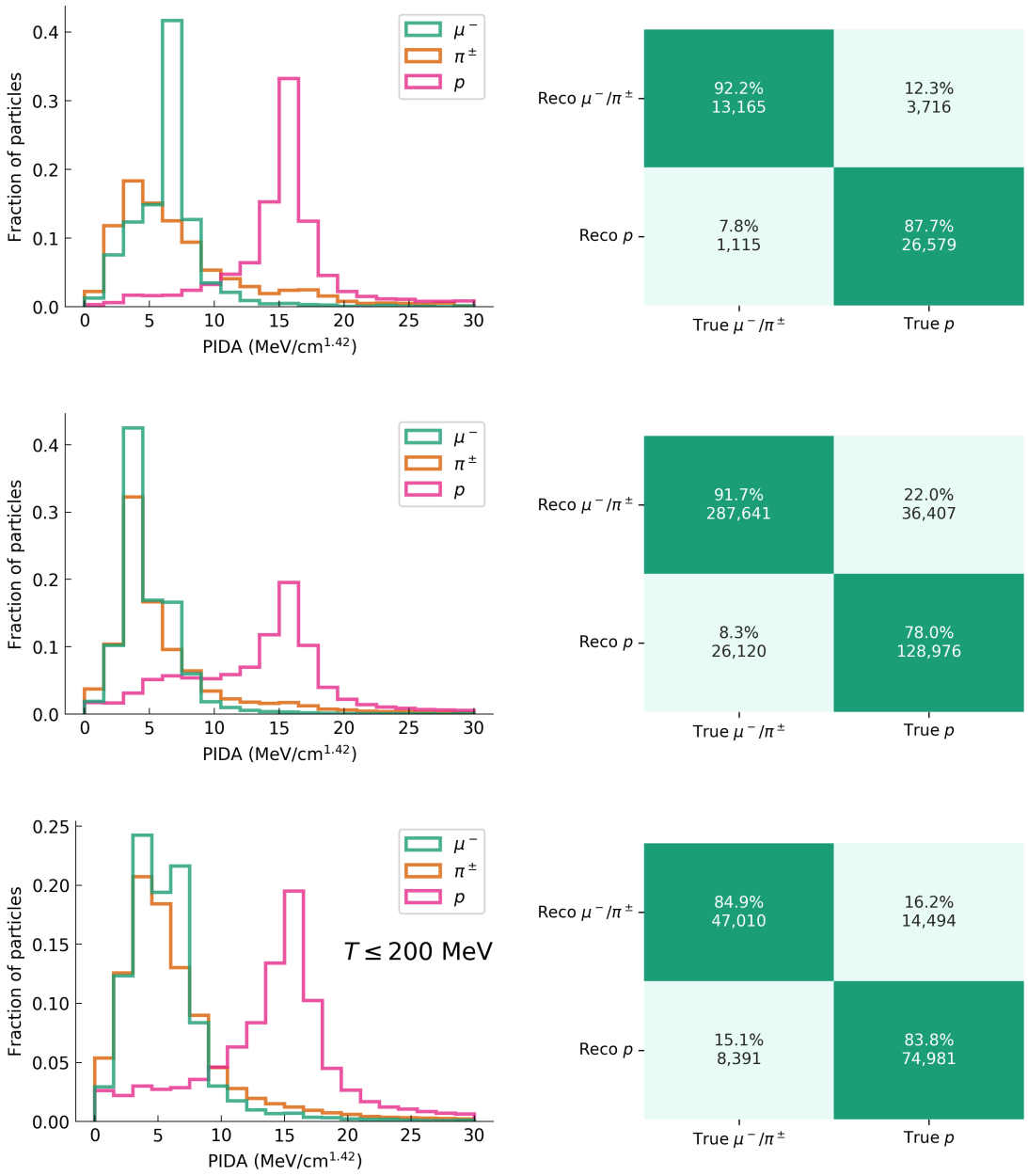


Figure 7.8: Performance of PIDA on the binary classification task distinguishing MIPs ($\mu^-/\pi^\pm$) from protons on the `MC_CLEAN_TRACKS` set (top row), `MC_RECO_TRACKS` set (middle row) and `MC_RECO_TRACKS_LE` set (bottom row), based on MicroBooNE simulation. The left column shows the distribution of PIDA for muons, charged pions and protons. The right column shows confusion matrices for 1D logistic regression classifiers trained to distinguish MIPs from protons using PIDA in each case. Accuracies as a function of particle energy are given in the FOMA comparison plots, Figures 7.10, 7.11 and 7.12.

7.4 Measuring the $\mathrm{d}E/\mathrm{d}x$ mode

To implement methods based on the models derived in chapter 6, we must first estimate the mode of the observed $\mathrm{d}E/\mathrm{d}x$ distribution. Two common ways of doing this are kernel density estimation (KDE) or a histogram-based approach. Both require choosing extra parameters: the kernel function and e.g. bandwidth for KDE, or the bin width for the histogram method. The quality of the mode found is dependent on this choice. To avoid the introduction of model parameters that may necessitate tuning, we instead opt for using the *repeated median estimator*: a nonparametric, robust method for straight line fitting, designed to tolerate samples from skewed and heteroskedastic distributions.

In repeated median regression, we fit the data $\{x_i, y_i\}$ for $i = 1, 2, \ldots, N$ to the regression line $y = A + Bx$, where A and B are scalars, using the estimators

$$\hat{A} := \underset{i}{\text{median}} \underset{j \neq i}{\text{median}} \frac{x_j y_i - x_i y_j}{x_j - x_i}, \tag{7.19}$$

$$\hat{B} := \underset{i}{\text{median}} \underset{j \neq i}{\text{median}} \frac{y_j - y_i}{x_j - x_i}, \tag{7.20}$$

where the median over $j \neq i$ produces a set of size N, and the median over i reduces it to a single scalar. This method has a breakdown point of 50%; i.e. up to 50% of the samples can be 'noise' before the accuracy of the estimator begins to degrade, where 'noise' for us means samples from the upper tail of the $\mathrm{d}E/\mathrm{d}x$ distribution (see Chapter 6) or real noise from poor reconstruction or detector effects.

7.5 FOMA

Using the first-order approximation of the $\mathrm{d}Q/\mathrm{d}x$ model developed in Chapter 6, we can construct a method that is robust to unmodelled effects and inaccuracies in the parameters ξ' and χ. Rearranging Equation 6.44,

$$\frac{\rho\mathcal{E}}{B}\left[\exp\left(\frac{BW_{\text{ion}}}{\rho\mathcal{E}C}\frac{\mathrm{d}Q}{\mathrm{d}x}\right) - A\right] = M(m,l)\frac{1}{\sqrt{R-x}}, \tag{7.21}$$

where

$$\begin{aligned} M(m,l) &:= \frac{1}{2}\sqrt{\xi'\chi(l)mc^2} \\ &= \frac{1}{2}\sqrt{\xi' mc^2\left(\ln\frac{2m_e c^2 \xi' l}{I^2} + j\right)}. \end{aligned} \tag{7.22}$$

That is, a plot of the LHS of Equation 7.21 (which is $\mathrm{d}E/\mathrm{d}x$ under the modified box model of recombination; see Section 6.4.1) against $1/\sqrt{R-x}$ for a Bragg peak ($R-x <$ 10 cm, as per Section 6.3.2) should yield an approximately straight line passing through the origin with gradient $M(m,l)$. Using the dependence of M on mass m and the known l, we can then identify regions of a parameter space that correspond to different particles in the detector. We note that, in practice, including l as a dimension adds little discriminatory power but introduces the complexity of classifying/clustering in an additional dimension, so we ignore this dependence.

Following Section 7.4, we can estimate this gradient as $\hat{B}$ from the repeated median estimator, and we expect that the intercept estimator $\hat{A}$ should be close to 0. In practice, we find that the intercept frequently deviates from 0. This observation likely has two main causes: offset particle end positions due to finite hit sizes and/or poor reconstruction leading to a constant offset in the residual range $R-x$, and error caused by taking the first-order approximation. To account for this, we consider classification within a 2D parameter space comprising the estimated gradient and estimated intercept, in contrast with the 1D space used by PIDA. Some examples of this are given in Figure 7.9. We call this variable construction method FOMA (first-order modal approximation).

7.6 Results and discussion

To permit direct comparison with the PIDA results shown in Figure 7.8, we classify in a supervised way by linearly separating the two groups using logistic regression. In practice, however, the classification is to be performed by unsupervised clustering in 2D (using a method such as k-means clustering) and identification of the clusters with particle types based on their position in this space, by analogy with PIDA. This makes the method less sensitive to data/MC disagreement. The performance is illustrated and compared with PIDA on the `MC_CLEAN_TRACKS`, `MC_RECO_TRACKS` and `MC_RECO_TRACKS_LE` sets in Figures 7.10, 7.11 and 7.12, respectively. The comparison is summarized in Table 7.1. Despite not relying on tuning using MC information, FOMA performance generally outperforms PIDA, with an AUC of 0.915 on the `MC_RECO_TRACKS` set compared with 0.887 for PIDA. The performance gain is most significant in the `MC_RECO_TRACKS_LE` set, corresponding to lower energy particles. This could be due to poorer reconstruction

Table 7.1: Comparison of the PID methods PIDA and FOMA on the MIP/proton classification task. The table shows the baseline accuracy, and the accuracy and AUC attained by the classifier. A classifier that simply chooses the most common particle type in the test set has an expected accuracy equal to the baseline accuracy and an expected AUC of 0.5. AUC is less sensitive to class imbalance so we use this statistic to identify the best classifier, which is displayed in bold. FOMA modestly outperforms PIDA on all test sets. Small variations in the baseline accuracies on the same test set are possible due to failed jobs, wherein only a random subset of the full test set was processed.

Test set	Method	Baseline acc. (%)	Accuracy (%)	AUC
MC_CLEAN_TRACKS	PIDA	68.0	89.2	0.917
	FOMA	68.0	90.8	**0.947**
MC_RECO_TRACKS	PIDA	65.5	87.0	0.887
	FOMA	65.5	88.5	**0.915**
MC_RECO_TRACKS_LE	PIDA	61.8	84.2	0.853
	FOMA	61.8	86.7	**0.901**

quality for low-energy events leading to anomalous hits; PIDA's use of the mean makes it susceptible to outliers, where FOMA's repeated median estimator affords it greater resilience.

7.6.1 Application to data

Since claims of decoupling from data/MC disagreement are features of both PIDA and FOMA, in this section we verify this on real data at MicroBooNE by comparing distributions on simulation and data. In MC simulated events, every event contains exactly one true neutrino, whereas only around 1 in 600 real BNB spills lead to a recorded neutrino interaction at MicroBooNE. This is partially corrected by the software trigger, which filters events based on PMT activity. To facilitate a fair comparison of MC and data distributions, we must further correct the distribution of simulated events to match the data. First consider the following datasets:

- MC_BNB_COSMIC: The set of simulated BNB events, including simulated cosmic background, using GENIE to generate neutrinos and CORSIKA to generate cosmic rays. Each event in this set corresponds to exactly one neutrino interaction.

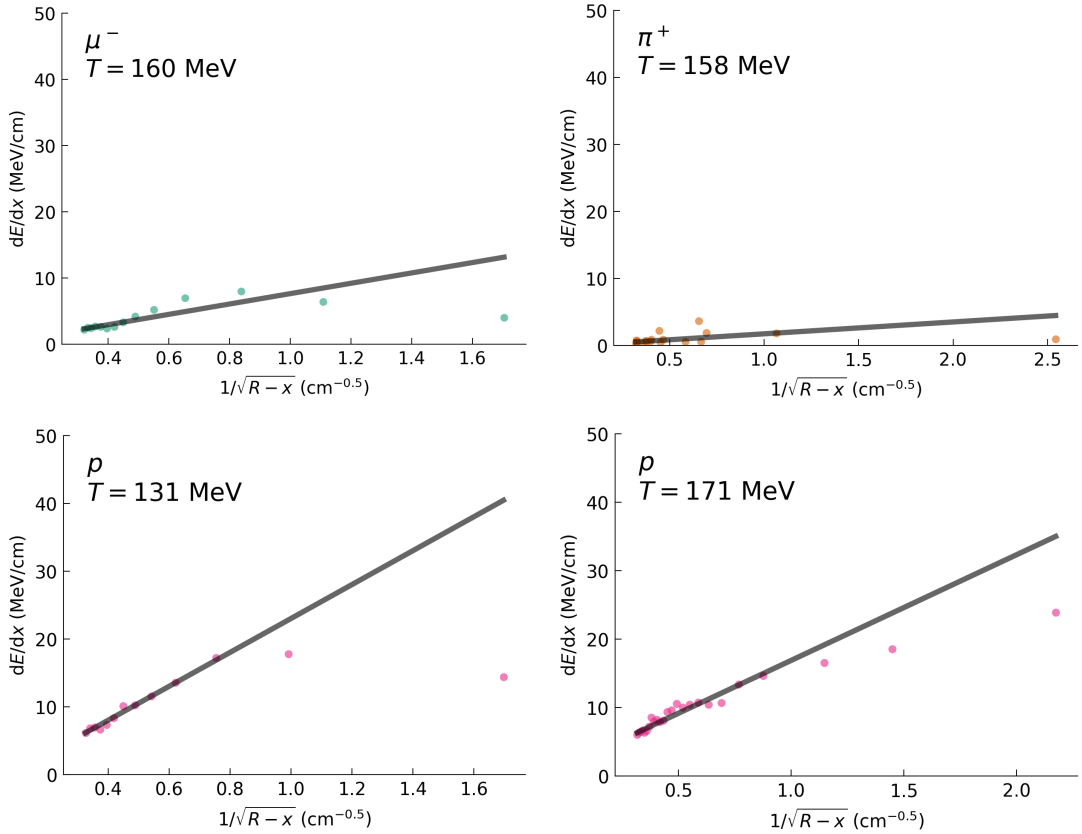


Figure 7.9: FOMA construction for hits in a random selection of muons, pions and protons. The line of best fit is given by the repeated median estimator, demonstrating the robustness of the fit. The leftmost hit in each plot corresponds to first hit within 10 cm of the end of the track, and the rightmost hit the last hit in the track. The greatest deviation from a straight line appears at the end of the track, where we would expect the first-order approximation to hold strongest. This could be due to errors stemming from the recombination correction at high $\mathrm{d}E/\mathrm{d}x$ and/or reconstruction effects at track boundaries.

- `BNB_Run1`: A real data sample taken during Run 1 at MicroBooNE. See Chapter 3 for a chronology of the runs. Each event corresponds to one BNB spill and is subject to the software trigger. A minority of events contains a neutrino interaction, so most events record only the cosmic background.

- `EXTBNB_Run1`: A real data sample taken during Run 1 at MicroBooNE, during a period when the BNB was switched off. Each event therefore records only cosmic background, with no BNB neutrinos.

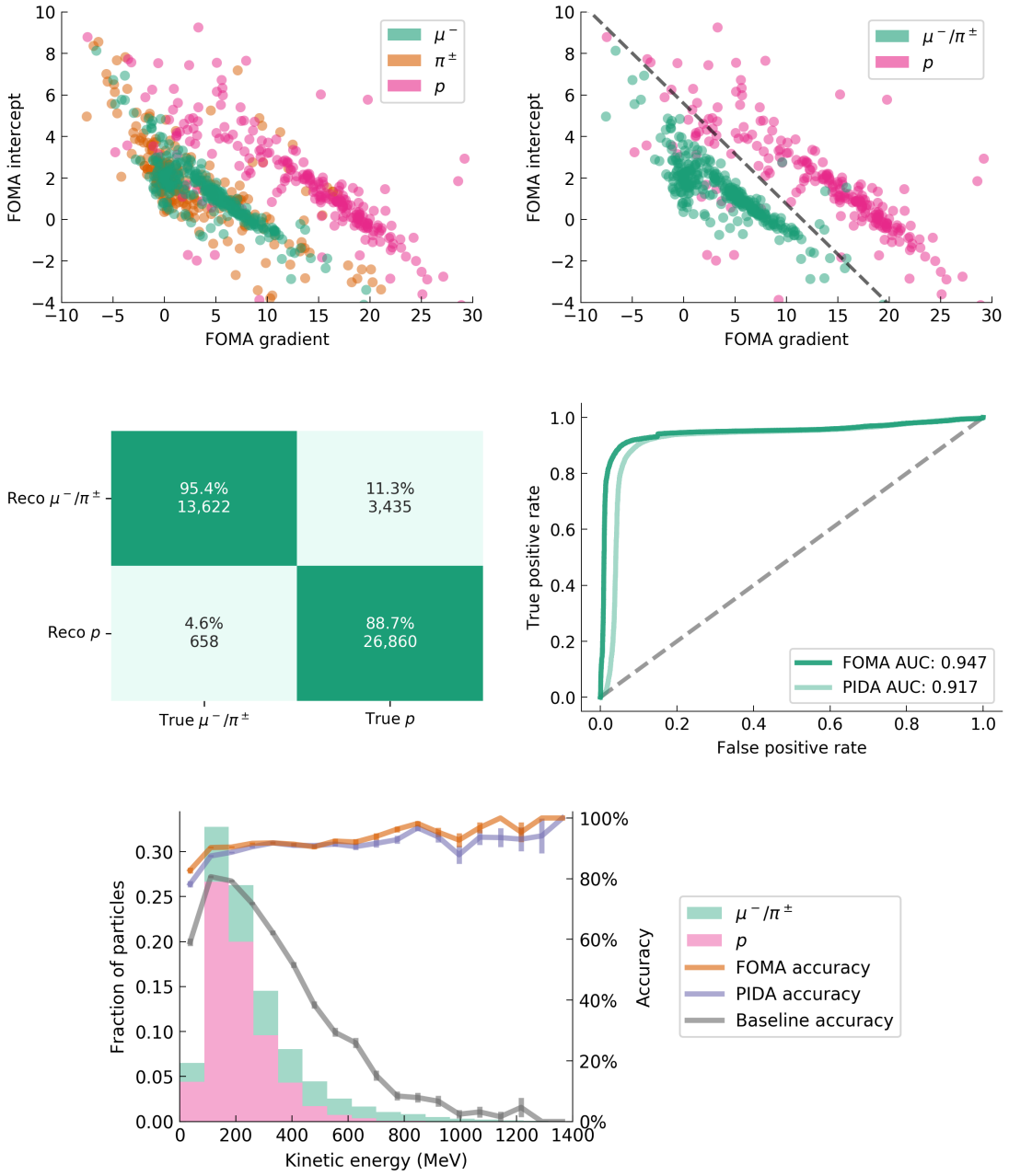


Figure 7.10: Performance of FOMA on the binary classification task distinguishing MIPs ($\mu^-/\pi^\pm$) from protons on the `MC_CLEAN_TRACKS` set. The cut shown in the top right plot is tuned on the training set via 2D logistic regression and performance on the test set is shown here. The accuracy baseline in the bottom plot is defined as the classifier that always chooses the most common particle type (here, protons).

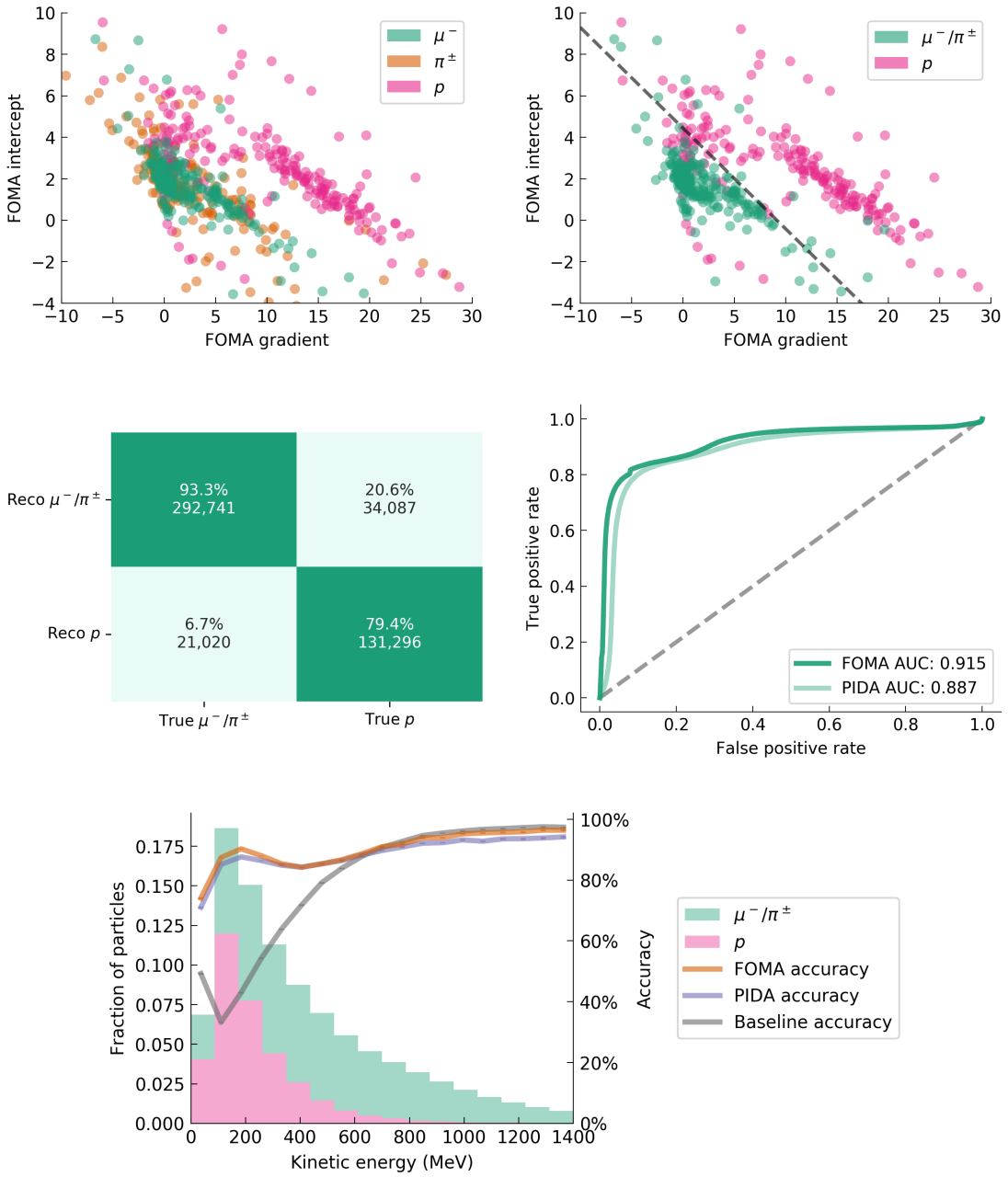


Figure 7.11: Performance of FOMA on the binary classification task distinguishing MIPs ($\mu^-/\pi^\pm$) from protons on the `MC_RECO_TRACKS` set. The cut shown in the top right plot is tuned on the training set via 2D logistic regression and performance on the test set is shown here. The accuracy baseline in the bottom plot is defined as the classifier that always chooses the most common particle type (here, muons).

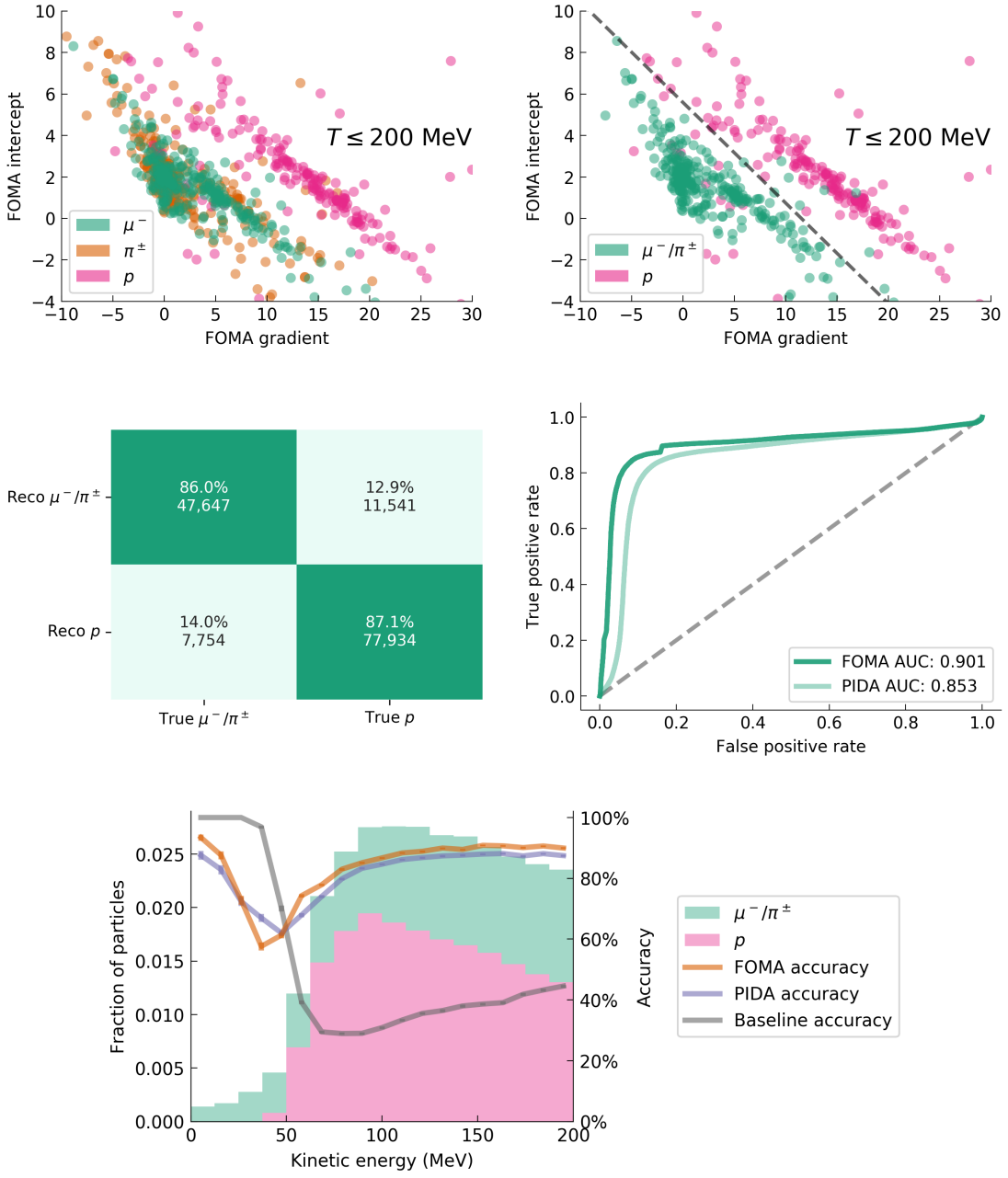


Figure 7.12: Performance of FOMA on the binary classification task distinguishing MIPs ($\mu^-/\pi^\pm$) from protons on the `MC_RECO_TRACKS_LE` set ($T < 200\,\text{MeV}$). The cut shown in the top right plot is tuned on the training set via 2D logistic regression and performance on the test set is shown here. The accuracy baseline in the bottom plot is defined as the classifier that always chooses the most common particle type (here, muons).

The procedure for comparing data with simulation for a given analysis follows the procedure laid out in [109] and is as follows:

1. Scale charge values in the `MC_BNB_COSMIC` and `BNB_Run1` samples by $4.113/5.076 \approx 0.810$. This accounts for the ADC-to-electron calibration factor being $(5.076 \pm 0.001) \times 10^3$ in simulation and $(4.113 \pm 0.011) \times 10^3$ in data [136].

2. Run the analysis over the `MC_BNB_COSMIC`, `BNB_Run1` and `EXTBNB_Run1` datasets, producing distributions for the variables of interest.

3. Rescale the `MC_BNB_COSMIC` distributions by $\alpha := \mathrm{POT}_{\texttt{BNB\_Run1}}/\mathrm{POT}_{\texttt{MC\_BNB\_COSMIC}}$, where $\mathrm{POT}_{\texttt{BNB\_Run1}}$ is the number of protons-on-target (POT) represented by the `BNB_Run1` set and $\mathrm{POT}_{\texttt{MC\_BNB\_COSMIC}}$ the number of POT represented by the `MC_BNB_COSMIC` set. This normalizes the exposure of the simulation to match that of the on-beam data sample.

4. Rescale the `EXTBNB_Run1` distributions by $\beta := \mathrm{TRIG}_{\texttt{BNB\_Run1}}/\mathrm{TRIG}_{\texttt{EXTBNB\_Run1}}$, where `BNB_Run1` is the number of BNB triggers and $\mathrm{TRIG}_{\texttt{EXTBNB\_Run1}}$ is the number of external triggers. This normalizes the number of spills/triggers of the off-beam data sample to match the on-beam data sample.

5. Rescale the `MC_BNB_COSMIC` distributions by $\gamma := (470/463.6)^2 \approx 1.028$. This corrects for a discrepancy between the simulated TPC z start position and the true position [143].

6. Add the scaled `MC_BNB_COSMIC` distributions to the scaled `EXTBNB_Run1` and compare these with the `BNB_Run1` distributions.

To isolate the case of MIP/proton separation, we wish to create high-purity samples of protons and MIPs, with completeness being less of a concern. We do this by introducing cuts on particle range R and angle to the beam direction ϕ. We consider fully-contained tracks that meet the `MC_RECO_TRACKS` cut described in Section 7.3, which can be applied to both data and simulation. This cut is dominated by the charged current (CC) $\mu + Np$ channels.

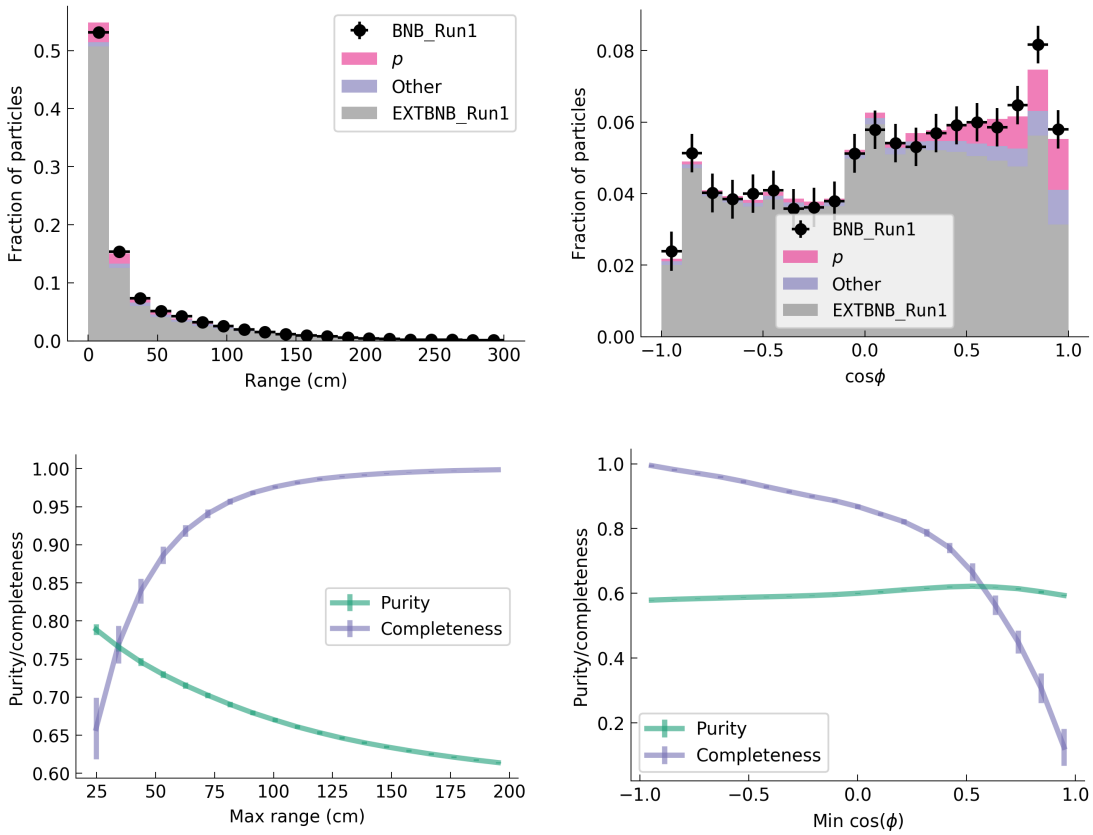


Figure 7.13: Selecting protons based on range R and beam opening angle ϕ. On the left hand side are unnormalised histograms showing the distributions of these parameters for protons and background. On the right hand side are shown purity and completeness plots for the proton selection taking each cut in isolation.

Proton selection

Figure 7.13 shows the distributions of R and $\cos\phi$ for protons and all other particles reconstructed as neutrino daughters in the BNB simulation, along with the purity and completeness resulting from maximum R and minimum $\cos\phi$ cuts taken in isolation. We roughly optimize them independently to get a high-purity sample without trading off too much completeness for practical purposes. This leads to cuts of $R_{\max} = 30\,\text{cm}$ and $\cos\phi_{\max} = 0.5$ (so $\phi_{\max} = 60°$) with a purity of 85.2% and a completeness of 13.3%. Purity and completeness are defined as

$$\text{purity}(\alpha) := \frac{|C(\alpha) \cap P|}{|C(\alpha)|},$$

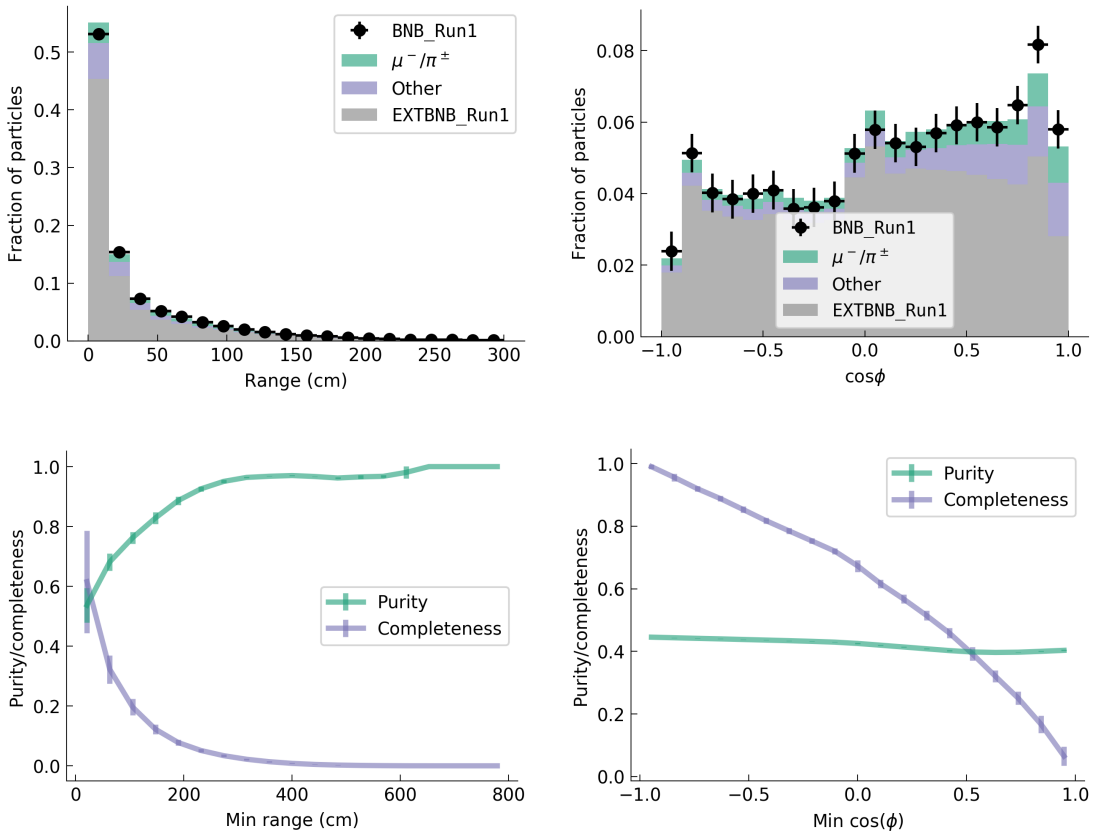


Figure 7.14: Selecting MIPs based on range R and beam opening angle ϕ. The top rows shows stacked histograms with the distributions of these parameters for MIPs and background. The bottom row shows purity and completeness plots for the muon selection taking each cut in isolation.

and

$$\text{completeness}(\alpha) := \frac{|C(\alpha) \cap P|}{|P|}$$

where

- $C(\alpha)$ is the set of particles passing the cut at threshold α, and
- P is the set of particles of interest that the cut is selecting for (here either the true set of protons or the true set of muons and charged pions).

The overlay shows reasonable agreement between data and simulation across the considered range and $\cos\phi$ spectra.

MIP selection

Figure 7.14 shows the distributions of R and $\cos\phi$ for MIPs and all other particles reconstructed as neutrino daughters in the BNB simulation, along with the purity and completeness resulting from minimum R and minimum $\cos\phi$ cuts taken in isolation. As for the proton selection, we roughly optimize them independently to get a high-purity sample without trading off too much completeness for practical purposes. This leads to cuts of $R_{\text{min}} = 300\,\text{cm}$ and $\cos\phi_{\text{max}} = 0.5$ (so $\phi_{\text{max}} = 60°$) with a purity of 99.9% and a completeness of 10.0%. Again we see reasonable data/simulation agreement across the range and $\cos\phi$ spectra.

Results of comparison

The PIDA and FOMA distributions for simulation and data are shown in Figures 7.15 and 7.16, respectively. PIDA offers easy comparison between data and simulation using a histogram to represent both the 1D PIDA value, whereas FOMA must be visualized in 2D to capture both variables. In both cases, the classifier response is 1D and can be studied using a histogram. Both PIDA and FOMA are able to distinguish between protons and muons in real data; this is evidenced by comparing the clear difference in distributions for the MIP cut and the proton cut in each case. In terms of a finer data/simulation agreement, the data peaks appear significantly smeared compared with simulation. There are a number of known reasons for data/MC disagreement in these samples:

- For tracks which are nearly orthogonal to wire planes, their ionization electrons induce signals that cause destructive interference on the induction planes and can lead to missing hits [119]. This effect, known as *dynamic-induced charged* (DIC), is unmodelled in simulation. In practice, the dependence of this unmodelled effect on track direction would lead to a smearing of distributions compared with simulation. This could help to explain the PIDA/FOMA peak smearing. This is now being modelled in MC simulation and taken into account in recent deconvolution work [119, 144].

- It has been demonstrated that CORSIKA underestimates the cosmic ray back-

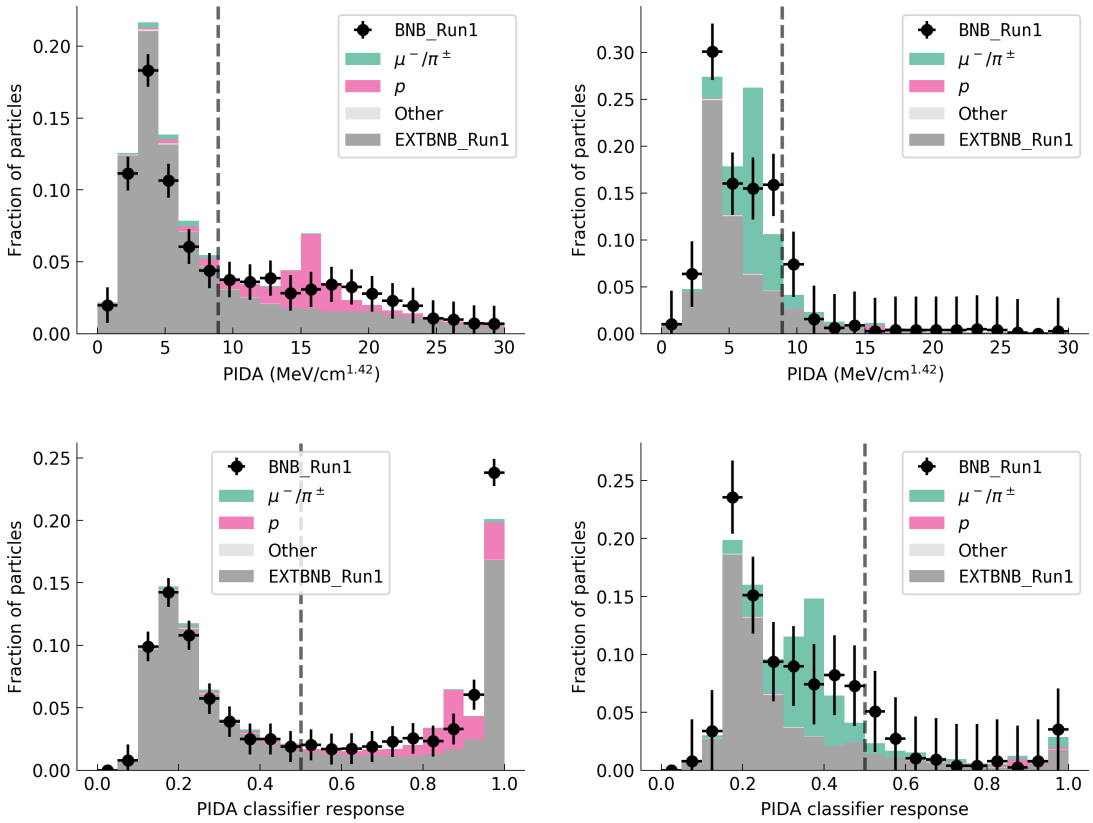


Figure 7.15: The distribution of PIDA for the proton cut (left column) and the muon cut (right column). The PIDA cut here is the one trained on the `MC_RECO_TRACKS` data subset illustrated in Figure 7.8. The bottom row shows the PIDA classifier response.

ground in the energy range $E < 100\,\text{MeV}$ [145]. Due to the minimum range cut that forms part of the MIP cuts, the scope of this effect is likely limited to the data/simulation balance in the proton plots.

- Simulated charged particle multiplicities have similarly been demonstrated to be overestimated by GENIE relative to data for the configuration used in this study. Events with greater charged particle multiplicities are usually more difficult to reconstruct due to a greater scope for hit merging and other reconstruction errors [127].

Overall, the data/simulation agreement is similar for PIDA and FOMA, which is not surprising given that they use of the same information (the 3D positions and charge

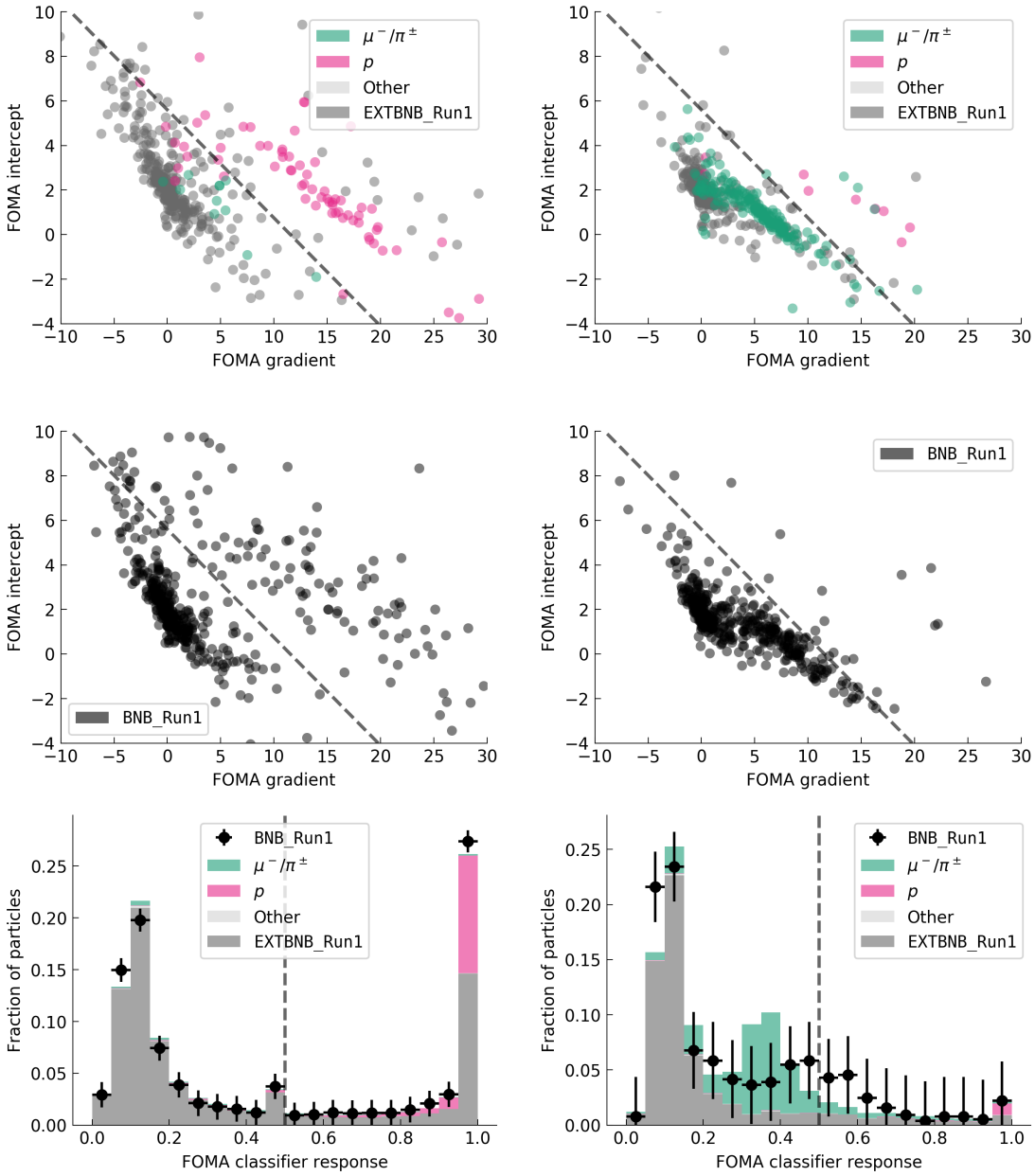


Figure 7.16: The distributions of the FOMA intercept and gradient for the proton cut (left column) and the muon cut (right column). The top row shows the distributions in simulation and the middle row in data. The cut here is the one trained on the `MC_RECO_TRACKS` data subset illustrated in Figure 7.11. The bottom row shows the FOMA classifier response.

depositions of the last $\sim 10\,\text{cm}$ of hits in each track). Finally, we note that the intended use of both PIDA and FOMA is not to use the MC cuts shown in these plots but instead to use the peaks/clusters observed in data to inform better cuts. The existence of peaks in PIDA and the existence of clusters in FOMA were indeed preserved, demonstrating that this procedure is possible. The peak smearing, however, could be a significant source of PID error. A promising improvement to PID analysis would be a method which can abstract a representation of the data that is less coupled to data/simulation disagreement, or actively learn to adapt based on real data.

Chapter 8

Representation learning for semi-supervised particle identification

This chapter proposes a novel semi-supervised particle identification (PID) method called PidNet based on representation learning. Following the domain adaptation philosophy of PIDA and FOMA, this method limits the extent to which data/simulation disagreement can affect the performance of a PID classifier trained on simulation but deployed on real data, in addition to facilitating interpretability. The $\mathrm{d}E/\mathrm{d}x$ profile of a particle is fed as a time-varying signal to a convolutional autoencoder, taking the latent state as an N-dimensional representation. By analogy with PIDA, the representation clusters can be labelled with particle types with reference with those produced by simulated (labelled) data. We demonstrate the superiority of this method to both PIDA and FOMA for distinguishing MIPs from protons.

8.1 Introduction

Development of the PIDA method can be considered an instance of *representation learning*: the $2N$-dimensional $\mathrm{d}Q/\mathrm{d}x$ profile of a track segment with N considered hits (comprising $(\mathrm{d}Q/\mathrm{d}x)_i$, x_i for $0 \leq i < N$) is mapped onto a one-dimensional variable. This

$\mathbb{R}^{2N} \to \mathbb{R}$ map is chosen such that different particles tend to occupy different regions in the new space. Whilst parameters of the map are tuned on simulated data, cut thresholds are chosen based on the distribution of the real (mapped) data, taking advantage of the fact that we can look at particles at a distribution level. Relying on simulated data only to define the map, rather than the cuts too, is an appealing solution as it partially protects against the impact of data/MC disagreement; such disagreement would result in a suboptimal map tuning for discrimination on real data. However, the method overall becomes less tightly coupled to distributional shifts.

In the case of PIDA, physical approximations are used to constrain the map optimization problem to two-dimensions, and the output space is chosen to be one-dimensional. The output space itself could be of any dimensionality: particles that occupy a higher-dimensional output space can analogously be grouped and labelled using unsupervised clustering techniques such as k-means clustering. A higher-dimensional output space allows richer information to be encoded but both the map optimization problem and downstream clustering tasks become more complex, entailing the possibility of greater error. A method able to generalize this concept to a higher number of dimensions requires adoption of modern techniques that can learn these higher-dimensional data representations.

A data-driven approach for learning these representations could take a number of forms, each with varying levels of parametrization. The first- and second-order methods developed in Chapter 7 represent an approximate but well-constrained approach driven by theoretical considerations, leading to one- and two-dimensional output spaces, respectively. Another approach for modelling sequential data is the *Kalman filter*, an algorithm for estimating the joint distribution of the variables of the hidden internal state of a linear stochastic dynamical system. It is underpinned by a hidden Markov model[1] (HMM) that assumes the state space of the latent variables is continuous and that all variables are normally distributed. Extensions that generalize Kalman filters to nonlinear systems include the extended Kalman filter and the unscented Kalman filter.

[1] A stochastic system based on a Markov chain, a sequence of events whose probability depends only on the state achieved by the previous event, which additionally contains unobservable states. HMMs have broad application to temporal pattern recognition problems such as speech recognition and handwriting recognition.

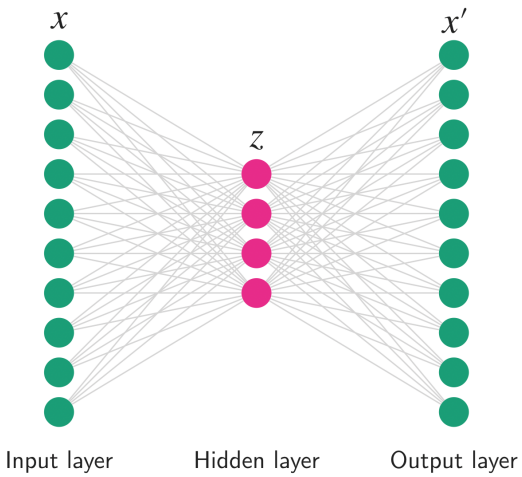


Figure 8.1: A toy example of an autoencoder, an unsupervised representation learning paradigm in which the network tries to learn to reconstruct its input despite a smaller hidden layer in the middle often called a bottleneck layer. This bottleneck layer can therefore be trained to encode an efficient compressed representation of the data z that preserves its salient features, such that $x \approx x'$. A real autoencoder typically has a more elaborate encoder and decoder both before and after the bottleneck layer.

These are used at MicroBooNE to assist with track reconstruction: before analysis, Pandora tracks are smoothed by Kalman filters passed forwards and backwards along the track. Assuming straight line propagation with some uncertainty, the algorithm iteratively takes into account 'new' information presented by each successive hit.

The use of machine learning and even deep learning in particle physics is not new [146–148]. Representation learning is an alternative to hand-crafting features that allows a model to automatically learn a feature space. This process can be supervised, self-supervised, semi-supervised or entirely unsupervised. In the supervised and semi-/self-supervised cases, internal representations are extracted from a network that is trained to perform one or more classification/regression tasks. In the unsupervised case, representations are extracted from an unsupervised model such as an autoencoder (see Figure 8.1). In this case, a fixed- and lower-dimensional encoding of the possibly time-dependent input is learnt without using labelled data.

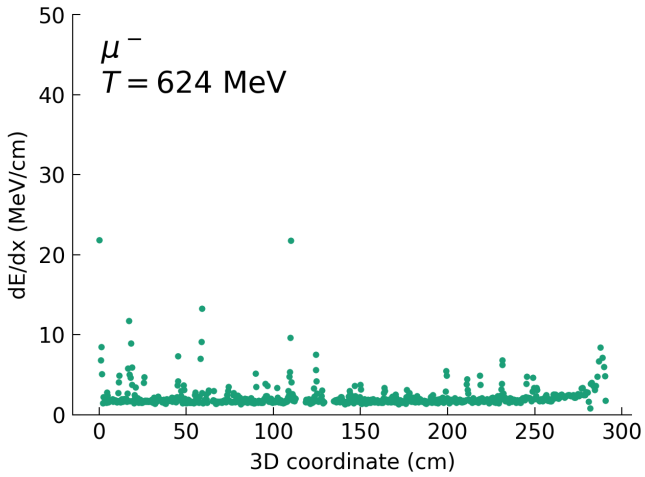


Figure 8.2: A visual representation of the calorimetric pair $\{(\mathrm{d}E/\mathrm{d}x)_i, \Delta x_i\}_{0 \leq i < N}$ for a simulated muon. Without explicitly including the hit width, the existence of detector gaps can only be inferred.

8.2 Particle identification as signal classification

In this chapter, we will cast the problem of particle identification (PID) for tracks as a multidimensional binary/multi-class signal classification problem. To make this casting, we first consider the calorimetric triplet described in Chapter 7: $\{(\mathrm{d}E/\mathrm{d}x)_i, \Delta x_i, w_i\}_{0 \leq i < N}$ for a track with N 3D hits. We ultimately wish to classify the particle type $y \in \{\mu^-, \pi^\pm, p, K^\pm, \ldots\}$ based on this information or a subset thereof. This problem is visually represented in Figure 7.5. Each component of the triplet can be considered a one-dimensional time-varying signal in which the 'time' dimension is the discrete hit sequence $i \in [0, N)$. The $\mathrm{d}E/\mathrm{d}x$ signal, even taken alone, has clear discriminatory power. We additionally include the Δx signal as it encodes relevant information about detector gaps and, indirectly, reconstruction quality and track longitudinality. To establish parity with PIDA/FOMA and limit complexity, we neglect the hit width and consider two-dimensional signal classification using $\{(\mathrm{d}E/\mathrm{d}x)_i, \Delta x_i\}_{0 \leq i < N}$. An illustration of the model input is illustrated in Figure 8.2.

Figure 7.5 illustrates the discriminative patterns a model could exploit, along with others which are irrelevant. Due to their similar masses, muons and pions are difficult to separate. We will therefore consider them together as MIPs, though the ability to

distinguish between them will be probed in detail. Sharper but smaller Braggs peak tend to suggest MIPs, whereas broader, higher peaks indicate protons. Even in the absence of a Bragg peak, the modal baseline for a proton is significantly higher than the $\sim 1.4\,\mathrm{MeV/cm}$ MIP level (see Chapter 6 for a full explanation). If a model is able to exploit track length[2], a shorter track increases the likelihood of a particle being a proton. Obvious but irrelevant features that should be ignored include detector gaps, poorly reconstructed/overlapping hits and high-$\mathrm{d}E/\mathrm{d}x$ 'noise' either from the tail of the Landau-Vavilov distribution (Chapter 6) or due to reconstruction pathologies.

A common problem in signal/image processing is a discrepancy between the distributions of real data and training data, leading to poor performance that often goes undetected. In particle physics, this can occur when a classification model is trained on simulated samples but deployed on real data that is subtly or significantly different in character. Tackling this problem requires adopting domain adaptation methods. PIDA and FOMA constitute two examples of this: the training data need not closely match the test data, as long as they are similar enough that the peak/cluster structure is largely preserved. Another method involves transforming the simulation so that it looks more like the data, which can be performed via calibration or using generative models [148].

8.3 Model architecture

8.3.1 Convolutional autoencoders

Deep learning can be employed for signal processing using a number of different neural network architectures, most commonly RNNs and convolutional neural networks (CNNs). CNNs are highly expressive models that deconstruct their input into hierarchical patterns, which are then progressively analyzed by each layer. This pattern of connectivity is inspired by the organization of neurons in the visual cortices of animals and, due to their translational invariance properties, they are most often applied to

[2]Architectural constraints may necessitate a fixed-length input, such that sequences shorter than this are padded and the model can therefore infer something about the track length. Whether this is desirable behaviour is arguable; this will be explored further in this chapter, along with ways of mitigating this dependence.

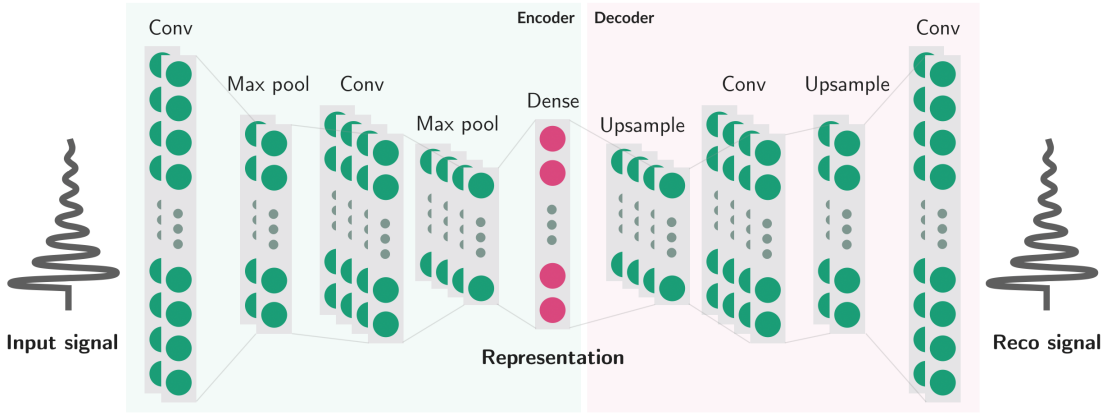


Figure 8.3: Example architecture for a one-dimensional convolutional autoencoder with two channels. The encoder comprises alternating convolutional layers and max pooling layers (see main text), culminating in a small, fully-connected ('dense') layer whose output is the vector representation. The decoder architecture mirrors the encoder and learns to reconstruct the input signal from this representation.

image analysis.

Each method presents limitations: RNN architectures suffer from the inability to learn long-range correlations [149], whereas CNNs can be computationally intensive. In this study, we opt for a CNN-based architecture to ensure that long sequences of hits can be analyzed with ease. Experimentation with alternative architectures is left to future work (Section 8.5.2). We combine a CNN with an autoencoder to make a convolutional autoencoder that can learn representations, as depicted in Figure 8.3. An autoencoder, once trained, takes a signal as input and produces a lower dimensional vector that preserves the most important information for reconstructing that signal; it can be considered a compression or dimensionality reduction method. The dimensions of the representation vector taken in isolation need not be physically meaningful, though it is possible to encourage this using disentanglement techniques [150, 151]. Such a model comprises an encoder and a decoder, which mirror one another in their structure so that the input and output dimensionalities are consistent. In the encoder, layers alternate convolutional layers, which consist of a set of filters that learn to activate when they detect a certain feature in their receptive field, and max pooling layers, which reduce dimensionality by replacing multiple non-overlapping regions across one or more dimensions with their maxima. The decoder substitutes max pooling layers for

upsampling layers, which duplicate each sample along the relevant axis/axes to undo the change in dimensionality caused by a pooling layer. For a more thorough description of CNNs, see [152].

8.3.2 PidNet architecture

The input to the autoencoder model, which we call PidNet, will be the calorimetric pair $\{(\mathrm{d}E/\mathrm{d}x)_i, \Delta x_i\}_{0 \leq i < N}$ following min-max scaling to assist training, i.e. two 1D signals. We therefore choose a 1D convolutional autoencoder with two channels. The details of the chosen architecture can be found in Appendix C. The architecture was inspired by those used in [153]. Since CNNs require fixed-size inputs, we pad/truncate the input hit sequence to be consistently 64 hits long.[3] Calorimetric pairs shorter than this are padded with zeroes. The sequence length is chosen to comprise enough hits to be able to characterize the Bragg peak, without becoming too complex to model or too computationally burdensome. The autoencoder is encouraged to encode the number of track hits if this number is less than 64, as it is important for accurate reconstruction of the input. Unlike PIDA and FOMA, PidNet is thus not necessarily invariant to particle length. This will be explored in Section 8.5. The dimensionality of the input space is $2 \times 64 = 128$. The size of the latent space (i.e. representation) is 16, so the network must learn to compress the information by a factor of 8. In practice, what information the network considers salient is guided by the loss function. In this case, we use binary cross entropy, which performs well for autoencoders with normalised input.

8.3.3 Classification

For every track, the autoencoder produces a 16-dimensional vector which can then be used for downstream tasks such as particle identification or energy estimation. We choose a standard neural network architecture for the final PID classification (see full architecture in Appendix C). The classifier is trained separately from the autoencoder in a supervised fashion on either the binary MIPs/proton classification task or the three-class muon/charged pion/proton classification task, using binary cross entropy

[3]The power of 2 facilitates a straightforward encoder/decoder mirroring that results in an output with the same dimensionality as the input.

or categorical cross entropy losses, respectively. The training and test sets are as in Chapter 7, with the exception that a random subsample of the training set has been selected as the validation set for optimization purposes.

An alternative to this two-step learning process (representation learning followed by classification) is to feed the signal straight into a CNN trained to perform the classification task, effectively replacing the decoder with the classification network. However, the two-step process offers greater robustness to MC fluctuations, the ability to train the autoencoder partially or entirely on real data, and greater scope for explainability (e.g. through qualitative interpretation of the representation space). In particular, the application of deep learning techniques to particle physics is complicated by the necessity of training on simulation and deploying on data, as the network is encouraged to exploit minute and potentially unrealistic features of the simulation to achieve the highest accuracy.[4] More concretely, we can imagine a network trained to perform PID on a dataset where $\mathrm{d}Q/\mathrm{d}x \in [1, 10]$ MeV/cm, which then makes predictions on data where $\mathrm{d}Q/\mathrm{d}x \in [0.7, 12]$ MeV/cm, related to the simulation through some stochastic nonlinear mapping. A deep learning classifier offers no guarantee that its response to such out-of-distribution samples varies smoothly or in any physically meaningful way, especially given the high dimensionality and noisiness of the input space. In the case of particle physics, these issues can also occur silently since we do not have access to true particle labels for real data. By performing a data-driven dimensionality reduction technique such as an autoencoder, the scope for this is drastically reduced since a well-trained autoencoder is forced to extract a small number of higher-level features (e.g. distribution shape parameters) due to the bottleneck layer. The resulting classifier is made robust by the smoother, lower-dimensional nature of its input, and data/simulation disagreement is straightforwardly studied by comparing representations at a distribution level. Furthermore, the learnt representations are agnostic to the task of PID and could be equally well suited to tasks such as data-driven energy estimation, particle directionality prediction and reconstruction error detection. The representation learning approach lays the groundwork for extensive future work, including actively addressing domain shift (see Section 8.5.2).

[4] This is known as overfitting.

8.4 Model training

The evolution of the representations can be tracked during training, both at a particle level and a distribution level. Figure 8.4 shows this evolution on a particle level for a proton. The reconstruction typically resembles a denoised version of the input signal, demonstrating the loss of information due to the bottleneck layer in the autoencoder. Figure 8.5 shows this on a distribution level. As the encoder learns, the representations become more able to distinguish between MIPs and protons, demonstrating that the representation is indeed preserving the information we need for PID, despite being unaware of the task. The representations are visualized using a t-stochastic neighbour embedding (t-SNE) plot [154], in this case a 2D projection of each 16-dimensional representation that tries to preserve the relationship between the data points. t-SNEs are stochastic and the axes are arbitrary, so each time one is created it may look superficially different. Furthermore, it is not generally possible to preserve all the information from 16 dimensions in 2 dimensions; the ability to identify clusters in a 2D t-SNE is therefore a sufficient but not necessary condition of being able to identify them in 16 dimensions.

8.4.1 Interpreting representations

To investigate the significance of clusters in the t-SNE space, we consider some correct and incorrect examples. The correct examples are given in Figure 8.6. The two more circular clusters correspond to (and separate) MIPs and protons that occupy the full 64 hits that the network can detect. The longer cluster represents particles with fewer than 64 hits, arranged along a sliding scale of particle range. While we did not intend the network to preserve this information, it has done so since it is useful for reconstructing the input.

The poor separation of MIPs and protons in the long cluster may be an artefact of the t-SNE projection and does not necessarily imply that they are inseparable in 16 dimensions. This hypothesis could be tested by investigating classification performance for particles with fewer than 64 hits. Range is a strong predictor of PID and therefore could unfairly boost the performance of this method, which is no longer merely

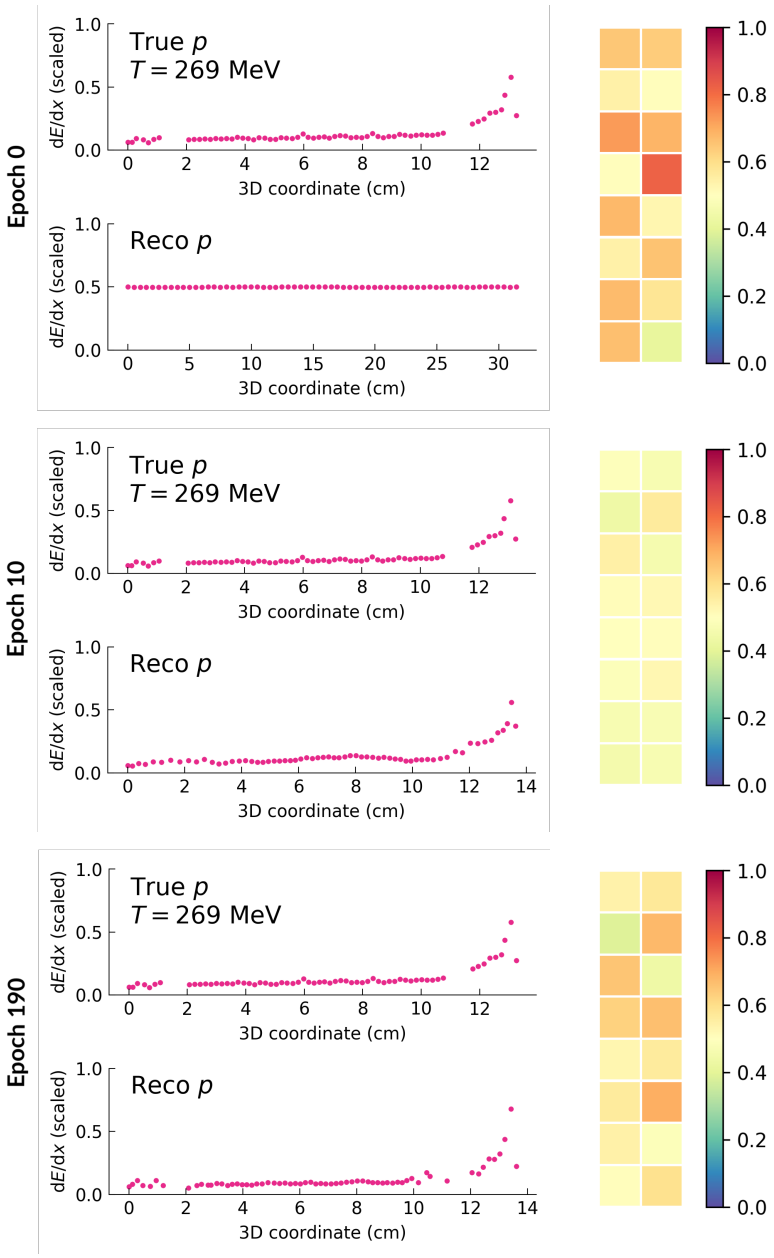


Figure 8.4: Evolution of a proton representation during training with snapshots at epochs 0, 10, 190 (top, middle and bottom rows, respectively). The preprocessed and reconstructed signals are shown in the left column along with a heatmap of the normalized values of the latent 16-dimensional representation in the right column. The heatmap represents the fixed-length encoding of the variable-length calorimetric signal. Its encoding is learnt by the encoder and its interpretation is learnt by the decoder/classifier during training. Its values are otherwise largely arbitrary, so its inclusion here is merely illustrative.

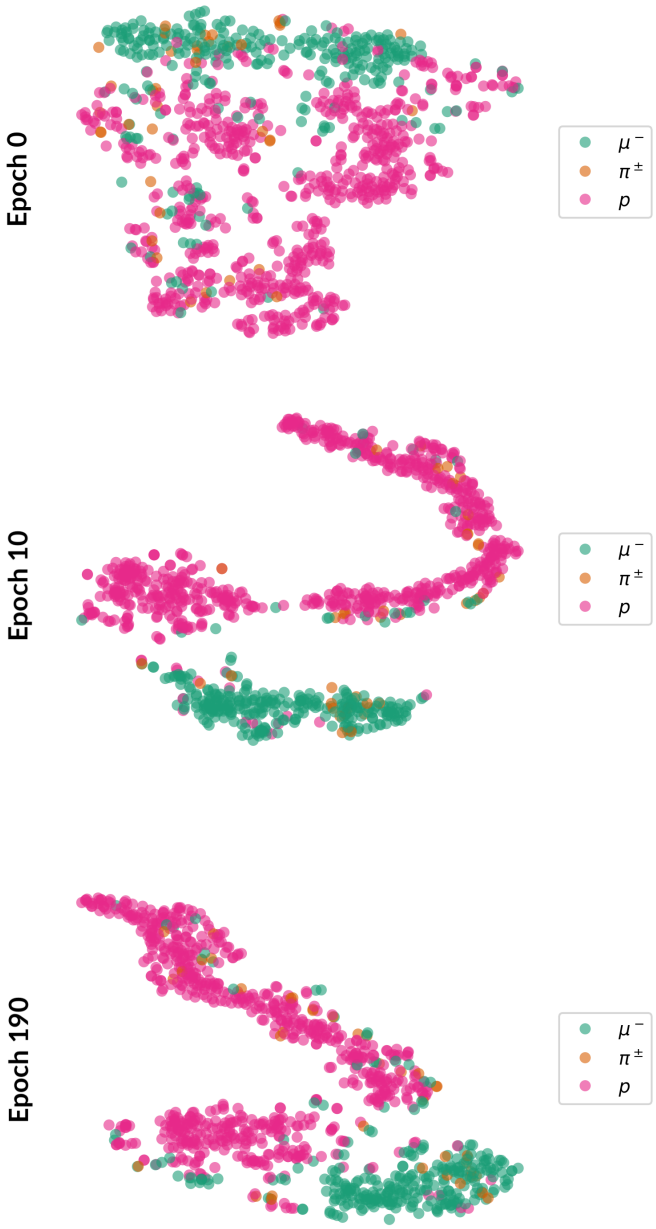


Figure 8.5: Evolution of the intermediate representations during training with snapshots at epochs 0, 10, 190 (top, middle and bottom rows, respectively). The 16-dimensional representations are represented here in two dimensions using a t-SNE (see main text).

characterising the shape of the Bragg peak. The claim that this method is superior to both PIDA and FOMA even in the absence of range information is substantiated in Section 8.5 and methods for constructing range-invariant representations are discussed in Section 8.5.2.

Examples of representations that lead to incorrect PID classification are shown in Figure 8.7. In these cases, the patterns of $\mathrm{d}E/\mathrm{d}x$ values near the Bragg peaks appear proton-like for the muon and MIP-like for the protons. This is a reassuring error and one that a human annotator could similarly have made in the absence of other information such as the nature of the other particles in the event. Such information could be taken into account in a downstream multivariate PID analysis.

8.5 Results and discussion

The validation procedure follows that of Chapter 7, including the definitions of the `MC_CLEAN_TRACKS`, `MC_RECO_TRACKS`, and `MC_RECO_TRACKS_LE` datasets. In this case, we also investigate whether we can distinguish muons from charged pions, something that was impossible for PIDA and FOMA. An example 2D t-SNE of the 16-dimensional representations fed into the classifier is shown in Figure 8.8 for both the two- and three-class cases.

PidNet performance on the `MC_CLEAN_TRACKS`, `MC_RECO_TRACKS` and `MC_RECO_TRACKS_LE` sets is examined in Figures 8.9, 8.10 and 8.11, respectively. The key results for the binary (MIP/proton) classification are summarized in Table 8.1 and the three-class (muon/charged pion/proton) classification in Table 8.2. The performance gain over both classical approaches is significant across almost the entire energy spectrum. Notably, we also achieve distinction between muons and charged pions with reasonable accuracy, though the subtle differences between these particles are likely to be blurred by data/simulation disagreement in practice. The performance gain is most evident on the `MC_RECO_TRACKS` set, where the AUC on the binary task is 0.979, compared with 0.914 for FOMA and 0.886 for PIDA. This performance gain can be explained through a number of factors:

- the network can easily approximate non-analytic functions such as an approxima-

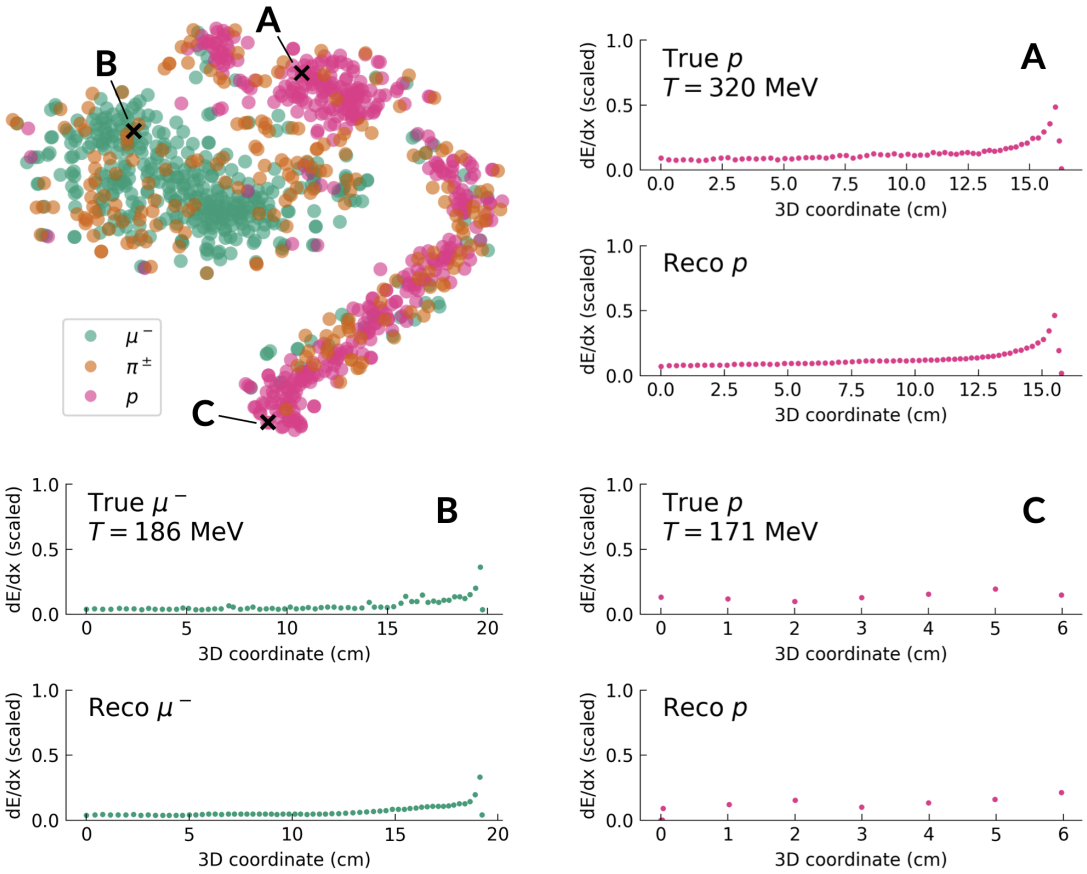


Figure 8.6: Examples of correctly-classified true signals and their reconstructions in different regions of the representation space. The top left plot shows an annotated 2D t-SNE of the 16-dimensional representation space. Three particles, indicated with black crosses, are selected and their true and reconstructed signals are displayed in the remaining three panels, correspondingly labelled as A, B or C.

tion to the Landau-Vavilov distribution;

- the network can learn to be robust to arbitrary noise, including common reconstruction errors and uncorrected detector effects such as gaps;
- the network has access to particle range for particles with less than 64 hits; and,
- the network does not need to assume the existence of a Bragg peak and may be able to classify non-stopping, backwards-going or poorly reconstructed particles using their $\mathrm{d}E/\mathrm{d}x$ distribution (see Chapter 6).

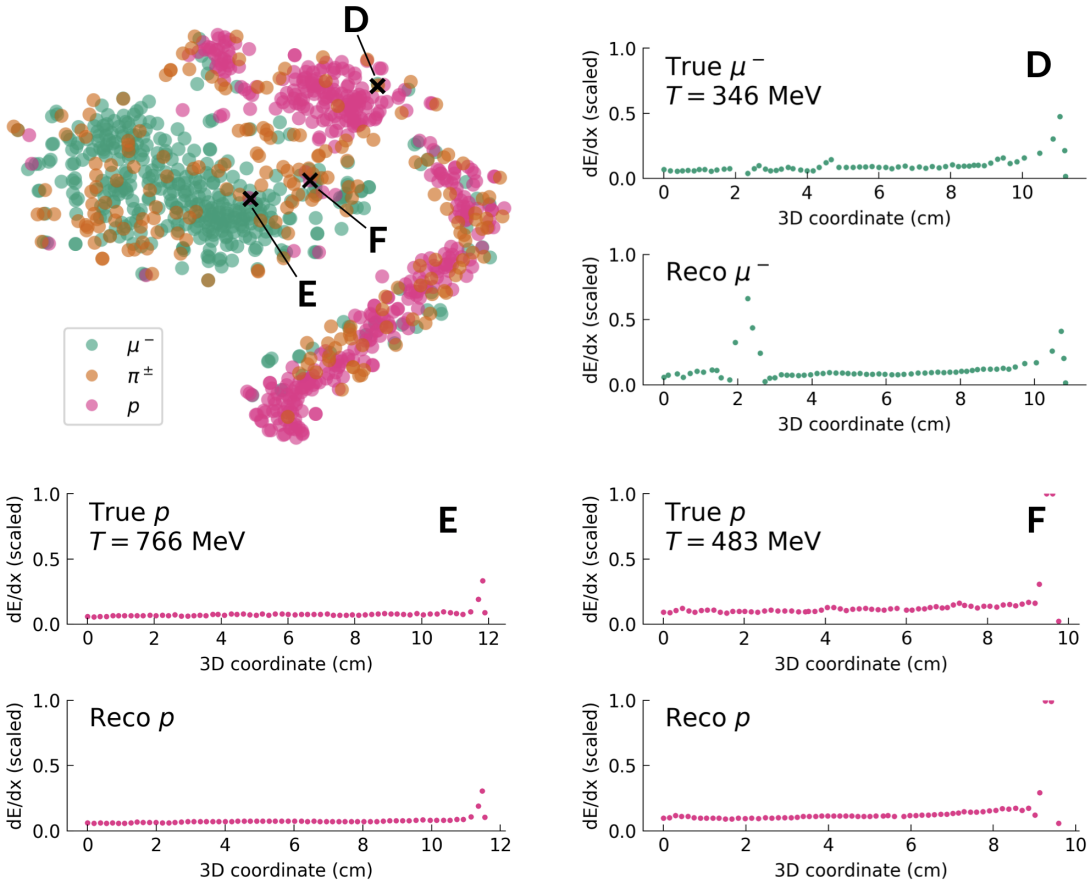


Figure 8.7: Examples of incorrectly-classified true signals and their reconstructions in different regions of the representation space. The top left plot shows an annotated 2D t-SNE of the 16-dimensional representation space. Three particles, indicated with black crosses, are selected and their true and reconstructed signals are displayed in the remaining three panels, correspondingly labelled as D, E or F.

All of these are desirable features, except for the particle range dependence. It would be desirable for a PID method to specialize in characterising the $\mathrm{d}E/\mathrm{d}x$ distribution alone, then particle range can be taken into account through a downstream multivariate analysis. This prevents a classifier[5] from learning a range prior for each particle type, e.g. that most short tracks are protons. While this might be a good rule for minimizing the loss function, it may perform poorly for certain kinds of analysis. To investigate

[5]Recall that the autoencoder itself does not know anything about the range prior for each particle type since it is trained in an unsupervised way, so this concern does not apply to the representations themselves.

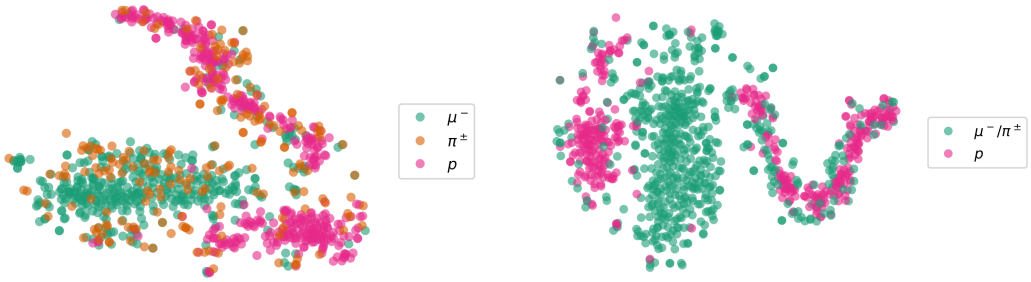


Figure 8.8: An example 2D t-SNE of the 16-dimensional representations fed into the downstream PID classifier for the three-class muon/charged pion/proton classification (left) and the binary MIP/proton classification (right). In this case, we represent a selection of particles from the `MC_RECO_TRACKS` set.

how much of the performance boost over PIDA/FOMA is due to range sensitivity, we study the performance of the classifier as a function of number of hits. This is shown in Figure 8.12 for the `MC_CLEAN_TRACKS` set and Figure 8.13 for the `MC_RECO_TRACKS` set. It is clear that, in both $N_{\text{hits}} < 64$ and $N_{\text{hits}} \geq 64$ regimes, PidNet maintains its superiority, suggesting that the range sensitivity does not play a major role.

8.5.1 Application to real data

Figure 8.14 illustrates the application of PidNet to real data using the same cuts and strategy employed in Section 7.6.1. The requirement that the number of track hits ≥ 64 simplifies the comparison and makes it fairer; for these particles, PidNet is invariant to particle length. These particles now predictably form two clusters without the long tail observed in Figure 8.8. As expected, the MIP cut consists of almost entirely MIPs and the proton cut a combination of MIPs and protons. To facilitate this comparison, we visualize representations using an isomap[6] instead of a t-SNE, which enables us to project new data into the space following the fit. In this case, we define the isomap space using MC particles that meet the non-range and -angle cuts, and then project the data representations into this space. The agreement between expectation and reality

[6]Both isomaps and t-SNEs are nonlinear dimensionality reduction techniques; isomaps are an extension of the isometric mapping method called multidimensional scaling (MDS), taking into account geodesic distances along the learnt manifold. Unlike t-SNE, new data can be projected onto the manifold after it has been fit.

Table 8.1: Comparison of the PID methods PIDA, FOMA and PidNet on the MIP/proton classification task. The table shows the baseline accuracy, and the accuracy and AUC attained by the classifier. A classifier that simply chooses the most common particle type in the test set has an expected accuracy equal to the baseline accuracy and an expected AUC of 0.5. AUC is less sensitive to class imbalance so we use this statistic to identify the best classifier, which is displayed in bold. PidNet significantly outperforms both PIDA and FOMA on all test sets. Small variations in the baseline accuracies on the same test set are due to failed jobs, wherein only a random subset of the full test set was processed.

Test set	Method	Baseline acc. (%)	Accuracy (%)	AUC
MC_CLEAN_TRACKS	PIDA	68.0	89.2	0.917
	FOMA	68.0	90.8	0.947
	PidNet	67.9	96.9	**0.993**
MC_RECO_TRACKS	PIDA	65.5	87.0	0.886
	FOMA	65.5	88.5	0.914
	PidNet	65.5	93.4	**0.979**
MC_RECO_TRACKS_LE	PIDA	61.8	84.2	0.850
	FOMA	61.8	86.7	0.899
	PidNet	61.8	90.4	**0.966**

Table 8.2: Performance of PidNet on the three-class ($\mu^-/\pi^\pm/p$) classification task. PIDA and FOMA are unable to distinguish between muons and pions so they are excluded from this study. The table shows the AUC for distinguishing each particle from all other particles, along with the overall baseline accuracy and classifier accuracy. Baseline accuracies and accuracies on a per-particle basis are omitted. A classifier that simply chooses the most common particle type in the test set has an expected overall accuracy equal to the overall baseline accuracy and all AUCs equal to 0.5.

	μ^-	$\pi^\pm$	p	Overall	
Test set	AUC	AUC	AUC	Baseline acc. (%)	Accuracy (%)
MC_CLEAN_TRACKS	0.979	0.879	0.990	67.9	86.6
MC_RECO_TRACKS	0.943	0.790	0.976	55.8	80.7
MC_RECO_TRACKS_LE	0.925	0.830	0.963	61.8	76.7

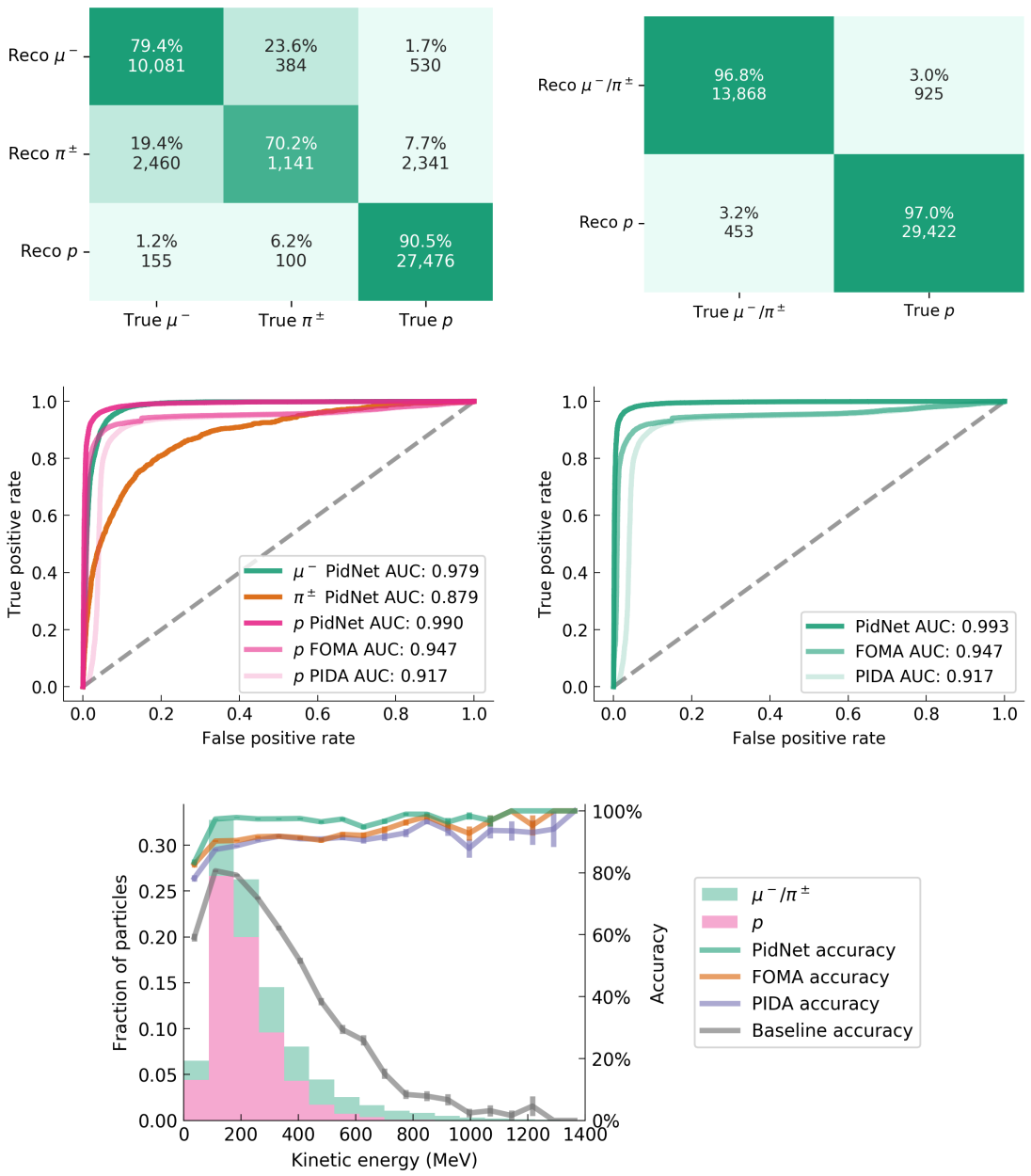


Figure 8.9: Performance of PidNet on the three-class classification task distinguishing muons, charged pions and protons, and binary classification task distinguishing MIPs ($\mu^-/\pi^\pm$) from protons on the `MC_CLEAN_TRACKS` set. The top row shows confusion matrices for the three-class (left) and binary (right) classification tasks. The middle row shows ROC curves for the three-class (left) and binary (right) classification tasks. The bottom row shows the accuracy as a function of particle kinetic energy; the baseline is defined as the classifier that always chooses the most common particle type (here, protons).

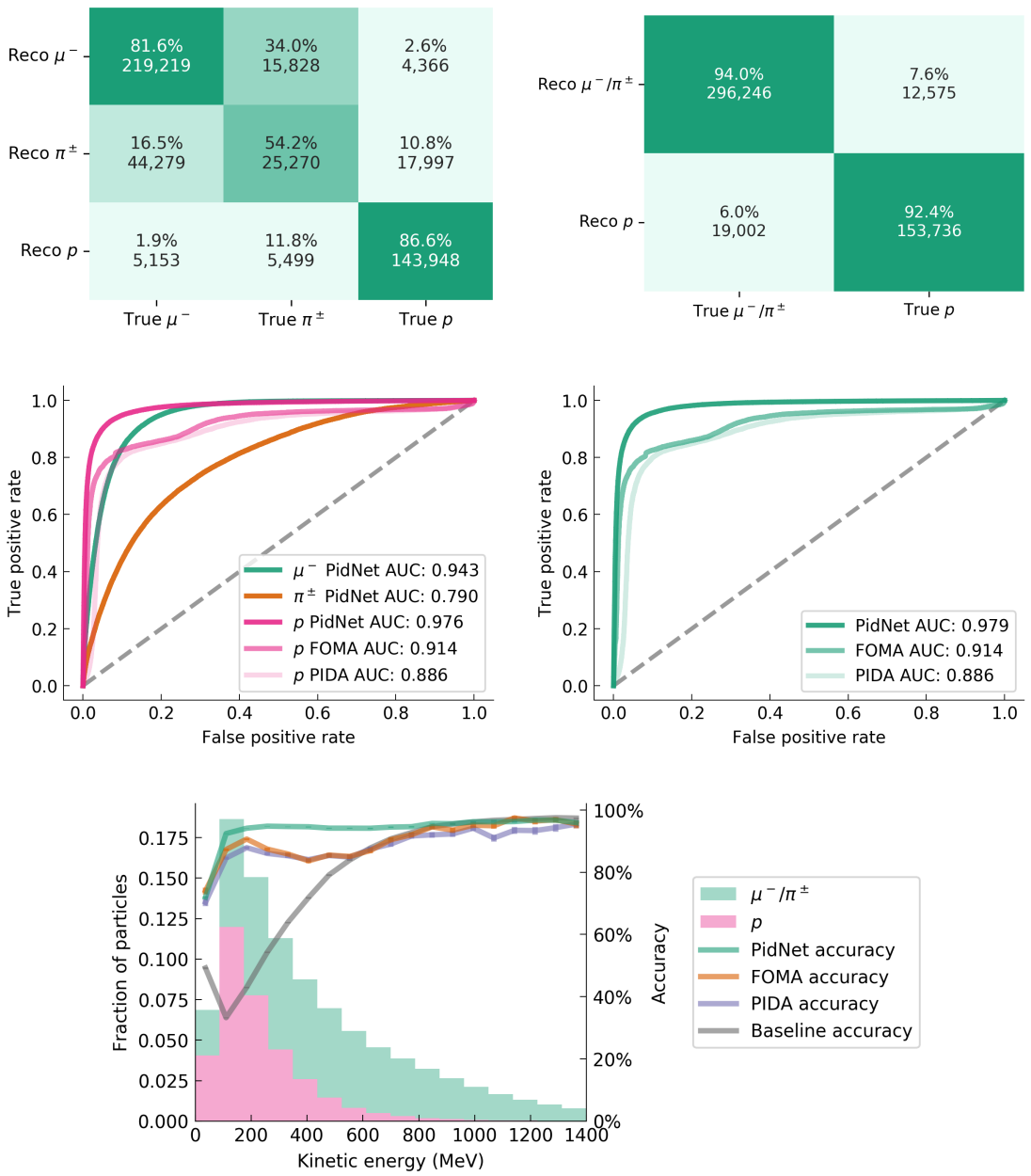


Figure 8.10: Performance of PidNet on the three-class classification task distinguishing muons, charged pions and protons, and binary classification task distinguishing MIPs ($\mu^-/\pi^\pm$) from protons on the `MC_RECO_TRACKS` set. The top row shows confusion matrices for the three-class (left) and binary (right) classification tasks. The middle row shows ROC curves for the three-class (left) and binary (right) classification tasks. The bottom row shows the accuracy as a function of particle kinetic energy; the baseline is defined as the classifier that always chooses the most common particle type (here, muons).

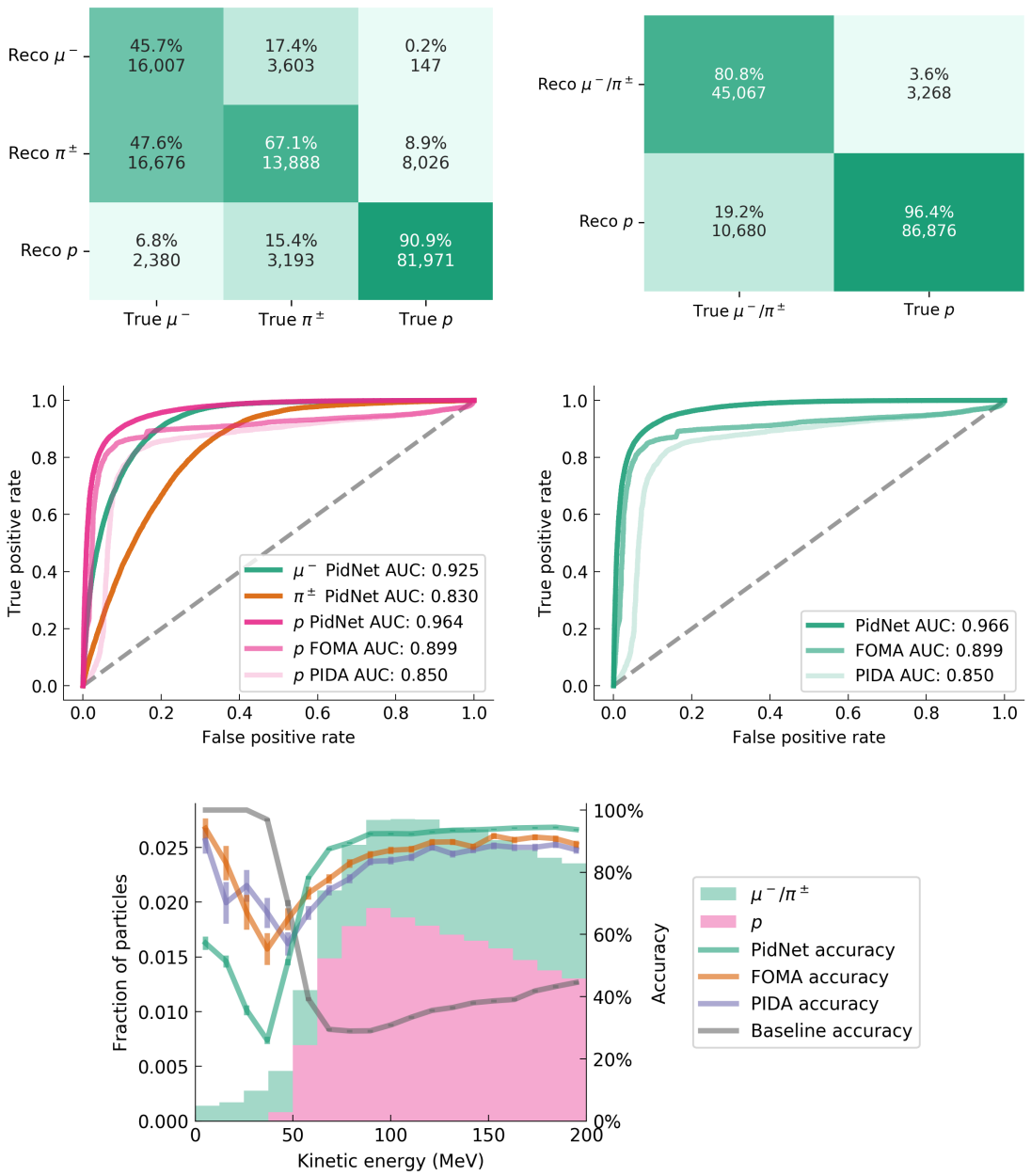


Figure 8.11: Performance of PidNet on the three-class classification task distinguishing muons, charged pions and protons, and binary classification task distinguishing MIPs ($\mu^-/\pi^\pm$) from protons on the `MC_RECO_TRACKS_LE` set ($T < 200\,\text{MeV}$). The top row shows confusion matrices for the three-class (left) and binary (right) classification tasks. The middle row shows ROC curves for the three-class (left) and binary (right) classification tasks. The bottom row shows the accuracy as a function of particle kinetic energy; the baseline is defined as the classifier that always chooses the most common particle type (here, muons).

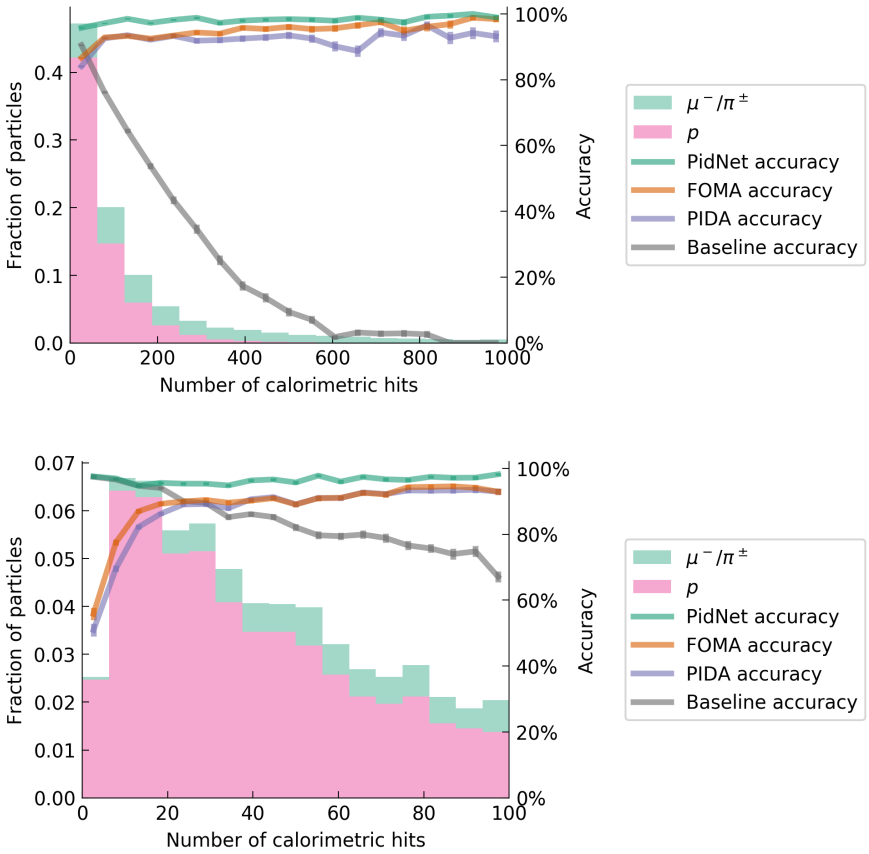


Figure 8.12: Performance of PidNet on the binary classification task distinguishing MIPs ($\mu^-/\pi^\pm$) from protons on the `MC_CLEAN_TRACKS` data set as a function of number of calorimetric hits. The top plot shows particles with numbers of hits in the range $[0, 1000]$; the bottom plot zooms into the range $[0, 100]$. The accuracy baseline is defined as the classifier that always chooses the most common particle type (here, protons). The superiority to PIDA and FOMA is maintained above and below the 64 hit threshold described in the main text.

in these plots gives us confidence that the information retained in the 16-dimensional representations is indeed meaningful and generalizable, rather than exploiting unrealistic aspects of the simulation. This is perhaps a surprising result as the expressiveness of the models gave great scope for overfitting to simulation. Requiring that the network compress this information into the low-dimensional representation has had the desired effect of retaining only the most salient information. If there had been more significant data/simulation disagreement, the performance of this method could suffer to a greater

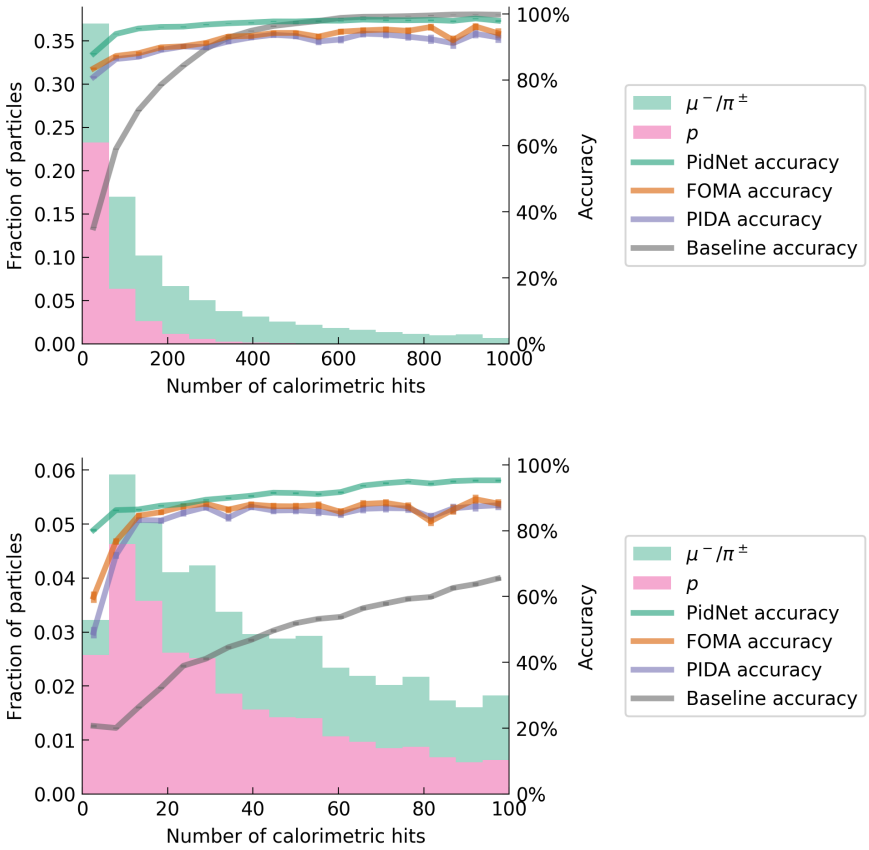


Figure 8.13: Performance of PidNet on the binary classification task distinguishing MIPs ($\mu^-/\pi^\pm$) from protons on the `MC_RECO_TRACKS` set as a function of number of hits. The top plot shows particles with numbers of hits in the range $[0, 1000]$; the bottom plot zooms into the range $[0, 100]$. The accuracy baseline is defined as the classifier that always chooses the most common particle type (here, muons). The superiority to PIDA and FOMA is maintained above and below the 64 hit threshold described in the main text.

extent. Future work that could solve this potential issue is discussed in Section 8.5.2.

8.5.2 Discussion

Improvements to the classification presented in this chapter can only be fairly minimal in terms of absolute performance gains, given the high accuracies demonstrated in the results. Regardless, there are a number of future developments that could improve its performance and/or generalizability.

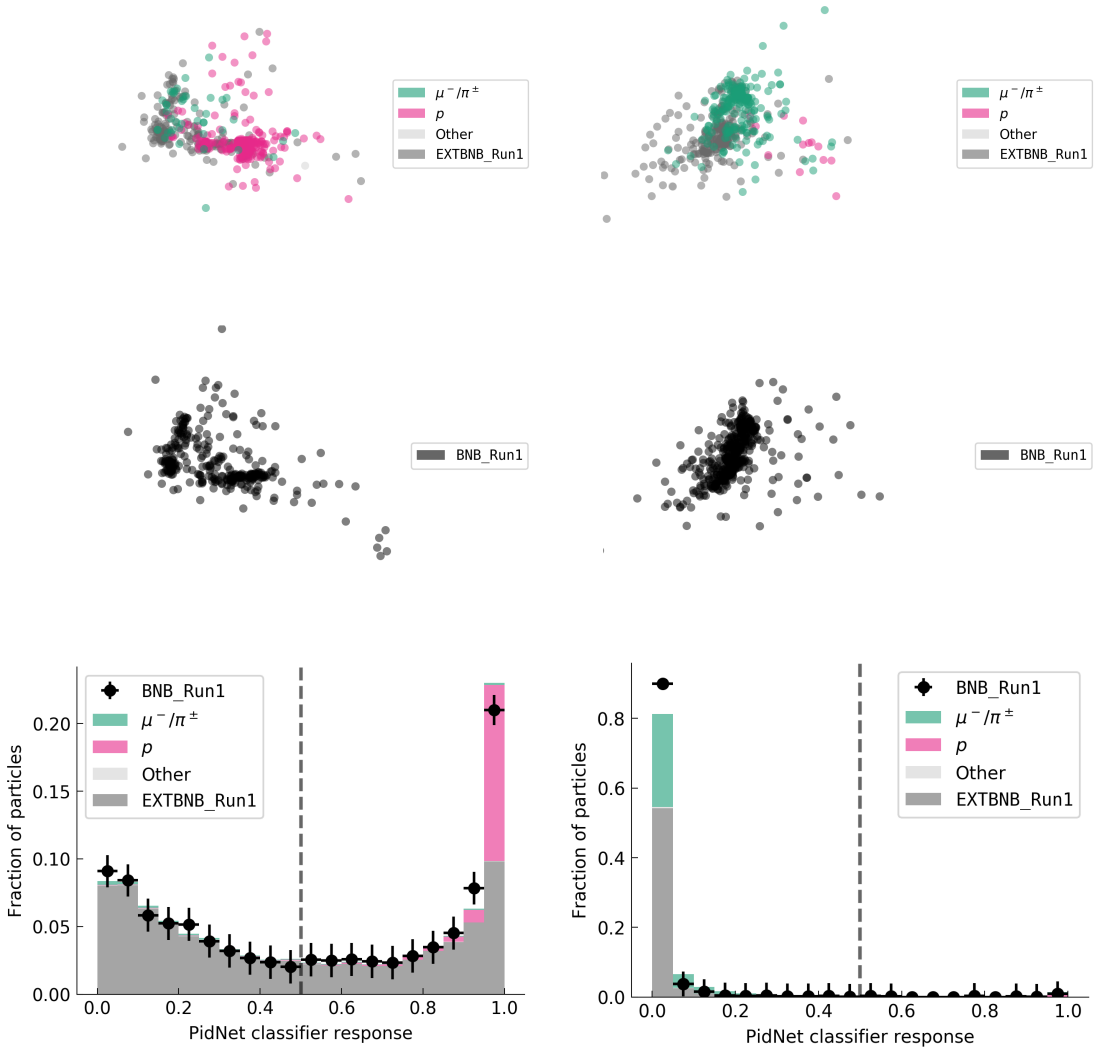


Figure 8.14: The application of PidNet to real data for the proton cut (left column) and the MIP cut (right column). At the top, simulation under the cuts described in Section 7.6.1 but without the R or $\cos\phi$ cuts is used to define an isomap space (coloured circles). The `MC+EXTBNB` sample is projected into this space (top row) along with the `BNB` sample (middle row). In all cases, we only consider tracks with at least 64 calorimetric hit reconstructed (see main text). The bottom row shows the PidNet classifier response.

Alternative autoencoder strategies

In this study, we considered the same autoencoder trained on the `MC_RECO_TRACKS` set and retrained the classifier for each new task. Since the autoencoder is trained in an unsupervised way, it could equally have been trained on real data, other MC simulation datasets or a combination thereof. This leads to two viable alternative strategies:

- training the autoencoder entirely on real data, then labelling the resulting representation clusters by performing inference on the MC simulation, optimizing the learnt representations for characterising data; or,
- training the autoencoder on a mixture of real data and MC simulation, then labelling the resulting data representation clusters by promixity to MC simulation clusters.

The latter is problematic as it is likely that any data/simulation disagreement will be exploited by the autoencoder to assist reconstruction, thus amplifying discrepancies. This could be mitigated through adversarial training, i.e. adding an additional output to the autoencoder that seeks to classify whether the example came from data or MC simulation, then backpropagating negative gradients.

Alternative architectures

The CNN autoencoder architecture employed for PidNet is a simple example in a vast landscape of possible architectures, and one that has presented issues relating to data/MC domain adaptation and particle range sensitivity. We can consider these two issues to be two sides of the same problem: the dependence of the learnt representations on undesirable covariates. Some architectures that may address these problems include:

- **LSTM/GRU autoencoders:** RNN-based transducers[7] such as these accept variable-length input and relatedly will not be encouraged to encode the sequence length, making them an obvious candidate for this problem. However, these tend to be poor at picking up on long-range correlations [149], which could be an issue

[7] That is, a sequence-to-sequence models that output one time step per input time step.

for long tracks. They also necessitate sequential computation (as opposed to the parallel computation possible with CNNs) making them less appealing.

- **Transformers:** The Transformer [155] architecture comprises chains of encoders and decoders using attention mechanisms and arranged in a particular way that permits modelling of long-range correlations and parallelization. Instances of Transformer-based architectures such as BERT [156] and GPT-2 [157] have replaced LSTM-based architectures as the state-of-the-art in a number of natural language processing tasks.

- **CPC:** Contrastive predictive coding [158] is an unsupervised representation learning paradigm that predicts the future in the latent space using autoregressive models and a probabilistic contrastive loss. CPC has been shown to be generally strong in tasks across speech, image, text and reinforcement learning.

- **VAEs:** Variational autoencoders [159] resemble normal autoencoders at a high level but are actually generative models that model their latent variables probabilistically to understand the underlying causal relations. The nature of their loss permits the injection of undesirable covariates into the latent space in such a way that encourages the learnt representation to be invariant to these parameters. This makes a compelling case for use in domain-adaptive and range-invariant PID.

Chapter 9

Impact of improved particle identification and vertex reconstruction on energy estimation

The impact on energy estimation of the improvements to neutrino interaction vertex reconstruction and PID is presented in this chapter. Methods for estimating the energy of showers and stopping tracks are described. A comparison of the vertex and PID algorithms demonstrates the improvements achieved by the new methods in this thesis. This comparison is performed using a nonparametric proxy for calorimetric energy resolution appropriate for skewed fractional energy difference distributions.

9.1 Introduction

Accurate measurement of the energy deposited in the TPC by the visible particles in the final state of the neutrino interaction is of great importance to MicroBooNE's physics goals. Indeed, investigating the low-energy excess of events originally observed at MiniBooNE relies directly on achieving high-quality track and shower energy reconstruction at low energies. The quality of this reconstruction is tightly coupled with a number of

factors:

- **Neutrino interaction vertex reconstruction:** The neutrino vertex is critical for accurate downstream reconstruction (see Chapter 4) and for accurate calorimetric profiling of particles, which is jeopardised if the vertex is placed e.g. in the middle of or at the wrong end of a particle. Improvements to this are developed in Chapter 5 in this thesis.

- **Particle range estimation:** Energy estimation based on particle range, given accurate PID, is typically robust. The algorithms developed in Chapter 7 to calculate the calorimetric triplet $(\mathrm{d}E/\mathrm{d}x, \Delta x, x)$ directly impact the estimation of the 3D particle range. Likewise, the theoretical groundwork for estimating energy from particle range in a LArTPC is laid out in Chapter 6.

- **Track/shower identification:** The method of particle energy reconstruction chosen depends on whether the particle is track-like, in which case we can calculate energy from range or by integrating the corrected and calibrated $\mathrm{d}E/\mathrm{d}x$ measurements along the track, or shower-like, in which case we sum, scale and correct the measured charge. Pandora uses a combination of geometric and calorimetric information to perform track/shower identification on 3D particles, as described in Chapter 4.

- **2D/3D hit reconstruction:** For both showers and tracks, the clustering of hits in 2D and their subsequent reconstruction in 3D is critical for reconstructing 3D paths and calorimetric information. The quality of this procedure is strongly affected by the accuracy of the neutrino vertex, which is significantly improved in Chapter 5.

- **Particle identification:** Calculating track energy from range requires accurate knowledge of the particle mass as well as the 3D path length. Development of new algorithms for identifying muons, pions and protons based on calorimetric information is the subject of Chapters 7 and 8.

All of these factors, except for track/shower identification, have been studied in the course of this thesis. We therefore wish to quantify the effect on the estimation of particle

energies of the two main subjects of this thesis: improvements to the neutrino interaction vertex reconstruction (detailed in Chapter 5) and the two new PID algorithms, PidNet and FOMA (detailed in Chapters 7 and 8).

In this chapter, we consider estimation of track and shower energies for fully contained, well-reconstructed particles with correct track/shower labels in order to help disentangle effects. For tracks, we use simulated charged current (CC) interactions from BNB ν_μ simulation with at least three reconstructed calorimetric hits; for showers, we use simulated CC interactions from BNB intrinsic ν_e simulation. The track validation is therefore dominated by the CC $\mu + p$ channel and the shower validation by the CC $e + p$ channel.

9.1.1 PID in practice

Estimation of track energy from range, which depends on the particle mass, is the preferred method in this chapter. Accurate distinction of MIPs and protons is therefore of great importance for the energy estimator. Chapters 7 and 8 compared and contrasted the performance of PIDA, FOMA and PidNet from the perspective of building binary/three-class classifiers in the idealized setting of only MIPs and protons. We wish to characterize our ability to accurately select MIPs and protons based on the classifier response. The results of this study for fully contained, well-reconstructed tracks from BNB simulation are shown in Figures 9.1, 9.2 and 9.3 for PIDA, FOMA and PidNet, respectively. Purity, efficiency and significance are defined as

$$\text{purity}(\alpha) := \frac{|C(\alpha) \cap P|}{|C(\alpha)|},$$

$$\text{efficiency}(\alpha) := \frac{|C(\alpha) \cap P|}{|P|}$$

and

$$\text{significance}(\alpha) := \frac{S(\alpha)}{\sqrt{S(\alpha) + B(\alpha)}},$$

where

- $C(\alpha)$ is the set of particles passing the cut at threshold α;
- P is the set of particles of interest that the cut is selecting for (here either the true set of protons or the true set of muons and charged pions);

- $S(\alpha) := |C(\alpha) \cap P|$ is the number of signal events at threshold α; and,
- $B(\alpha) := |C(\alpha) \setminus P|$ is the number of background events at threshold α.

In all cases, a cut at the default value of 0.5 achieves near-optimum performance. As expected from the results in Chapter 8, PidNet is able to select MIPs and protons considerably better than PIDA and FOMA. The shape of the classifier response is markedly different between the classifiers; PIDA and FOMA are explicitly trained to have probability-like outputs, whereas PidNet's output is encouraged to be polarized by its loss function.[1] Contributions from track-like particles that are not MIPs or protons are minimal. For the remainder of this chapter, PID is performed using these classifiers by thresholding the output at 0.5; an output above this threshold identifies protons, below identifies MIPs. All calorimetric work is performed using the collection plane alone, as justified in Chapter 7.

9.2 Estimating shower energy

Following the procedure developed by the ArgoNeuT collaboration, shower energy is estimated by converting the observed hit charge (e^-_{reco}) induced by electrons and photons into energy depositions and summing them [160]. The conversion to energy depositions requires multiplying the charge deposition by the following calibration constant [161]:

$$\frac{E^{\text{e}/\gamma}_{\text{reco}}(\text{MeV})}{e^-_{\text{reco}}} = 1.01 \frac{e^-}{e^-_{\text{reco}}} \times \frac{23.6\text{eV}}{e^-} \times 10^{-6} \frac{\text{MeV}}{\text{eV}} \times \frac{1}{R} = 3.85 \times 10^{-5}, \tag{9.1}$$

where

- e^- is the charge on an electron;
- $1.01 e^-/e^-_{\text{reco}}$ is an empirical factor that corrects the underestimation of the true number of electrons collected on the wires; its value is obtained by measuring the number of collected electrons using a sample of stopping muons and performing a fit to the $\text{d}E/\text{d}x$ profile expected from reference data [13];
- $23.6\,\text{eV}/e^-$ is the work function of liquid argon [162]; and,

[1] It is possible to encourage a neural network to have a more probabilistically meaningful output through Bayesian machine learning or, more recently, normalizing flows.

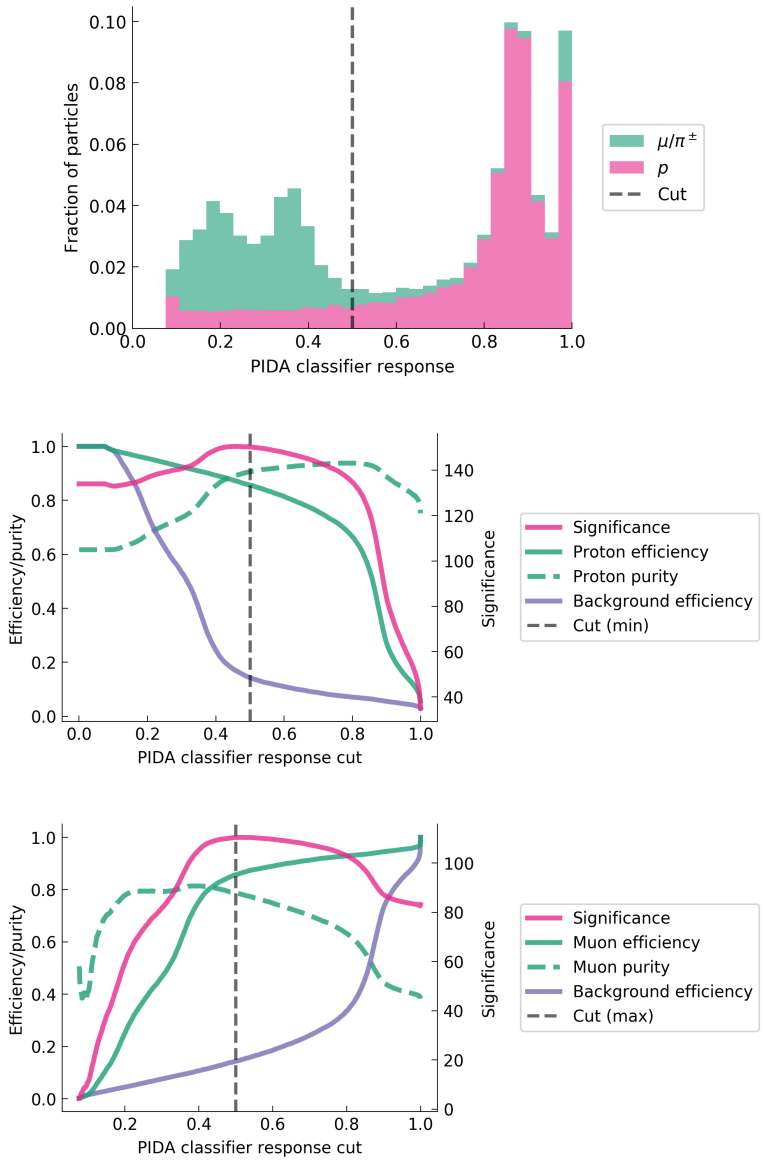


Figure 9.1: Performance of the PIDA classifier described and trained in Chapter 7 on well-reconstructed, fully contained tracks from BNB simulation. The PIDA classifier output, which is probability-like, is shown in the top plot as a stacked histogram for true MIPs and true protons. The efficiency, purity and significance of the cut for proton and muon selection is given in the middle and bottom plots, respectively. The default cut at 0.5 is indicated.

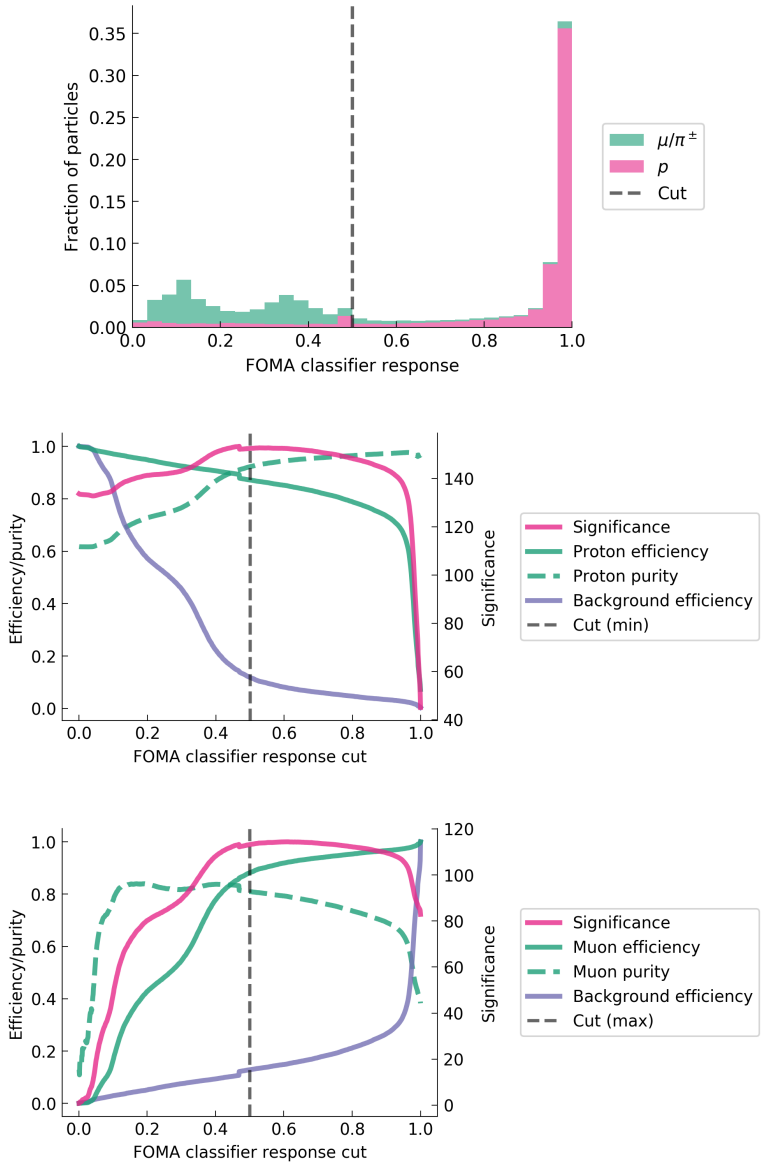


Figure 9.2: Performance of the FOMA classifier described and trained in Chapter 7 on well-reconstructed, fully contained tracks from BNB simulation. The FOMA classifier output, which is probability-like, is shown in the top plot as a stacked histogram for true MIPs and true protons. The efficiency, purity and significance of the cut for proton and muon selection is given in the middle and bottom plots, respectively. The default cut at 0.5 is indicated.

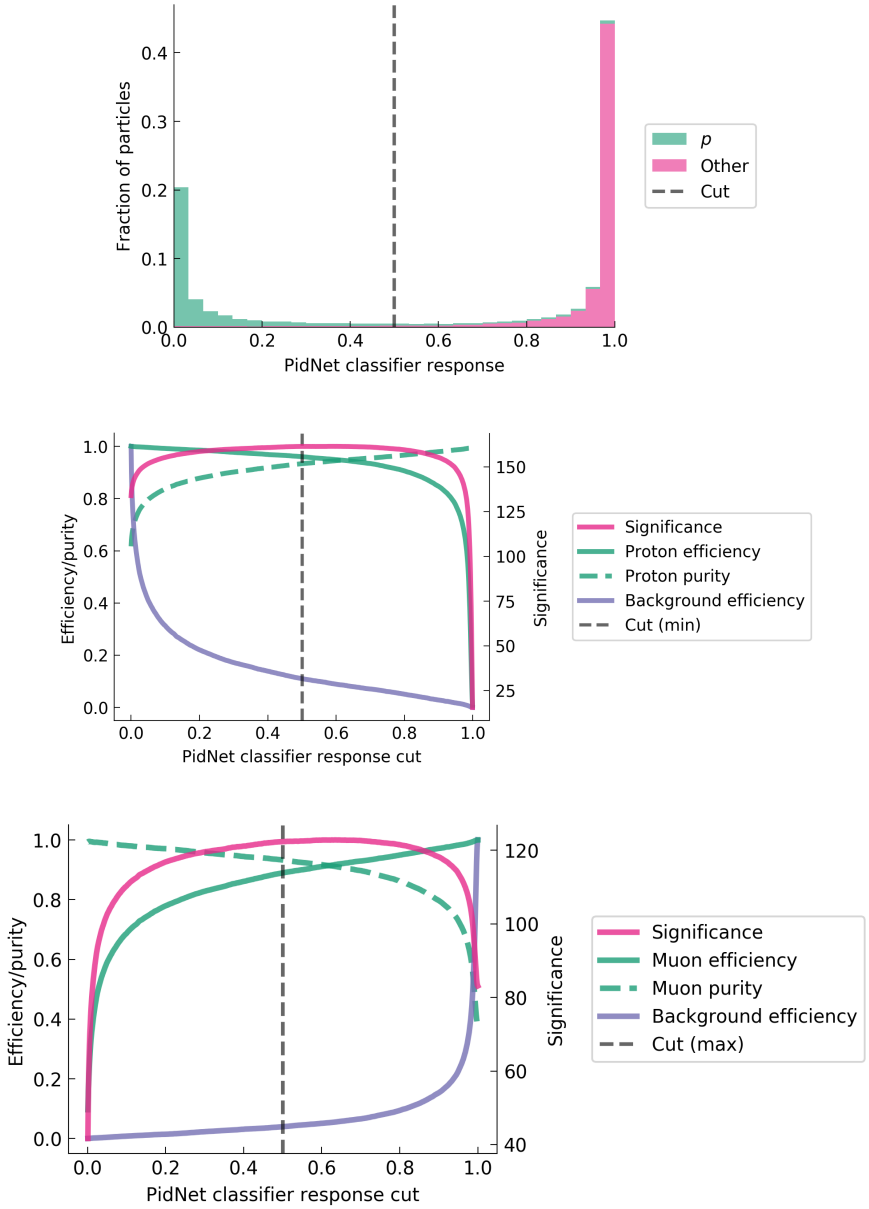


Figure 9.3: Performance of the PidNet classifier described and trained in Chapter 8 on well-reconstructed, fully contained tracks from BNB simulation. The PidNet classifier output, which is influenced by the shape of the loss function, is shown in the top plot as a stacked histogram for true MIPs and true protons. The efficiency, purity and significance of the cut for proton and muon selection is given in the middle and bottom plots, respectively. The default cut at 0.5 is indicated.

- $R = 0.62$ is the recombination correction for a MIP in a MicroBooNE-like detector using the Modified Box model (see Section 6.4.1) [140].

Linear calibration of the shower energy estimation is performed by dividing the true energy spectrum into 10 equally sized bins in the range 30 MeV to 2,030 MeV, chosen such that they are all well-populated, and comparing the observed (asymmetric) distribution with the bin centre, as shown in Figure 9.4. There are two common fits used to model lossy processes like this: the Crystal Ball function and its simpler variant, the GaussExp function [163]. The (unnormalized) Crystal Ball function is defined as

$$f_{\mathrm{CB}}(x;\alpha,n,\bar{x},\sigma) := \begin{cases} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{\alpha^2}{2}\right)\left(\frac{n}{|\alpha|} - |\alpha| - \frac{x-\bar{x}}{\sigma}\right)^{-n}, & \text{for } \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases}$$

where α, n, $\bar{x}$ and σ are fit parameters. The (unnormalized) GaussExp function is defined as

$$f_{\mathrm{GE}}(x;\bar{x},\sigma,k) := \begin{cases} \exp\left(\frac{(x-\bar{x})^2}{2\sigma^2}\right), & \text{for } \frac{x-\bar{x}}{\sigma} > -k \\ \exp\left(\frac{k^2}{2} + k\left(\frac{x-\bar{x}}{\sigma}\right)\right), & \text{for } \frac{x-\bar{x}}{\sigma} \leq -k \end{cases}$$

where $\bar{x}$, σ and k are fit parameters. The Crystal Ball function stitches together a power law distribution with a Gaussian, whereas GaussExp uses an exponential decay in lieu of the power law, which leads to more robust fitting procedure [164]. Having tested both, we opted for the Crystal Ball function, which is able to provide a significantly better fit to the data. The most likely value of this fit in each bin is used to create the calibration curve, which is shown in Figure 9.5, yielding

$$E^{e/\gamma}_{\mathrm{reco}} = 0.75 E^{e/\gamma}_{\mathrm{true}} - 0.21\,\mathrm{MeV}. \tag{9.2}$$

The calibrated shower energy is therefore given by

$$E^{e/\gamma}_{\mathrm{calib}} = \left(E^{e/\gamma}_{\mathrm{reco}} + 0.21\,\mathrm{MeV}\right)/0.75. \tag{9.3}$$

The calibrated shower energy measurements can then be used to estimate the calorimetric energy resolution as a function of the true shower energy. We define the fractional energy difference as

$$E_{\mathrm{frac}} := \frac{E_{\mathrm{calib}} - E_{\mathrm{true}}}{E_{\mathrm{true}}}. \tag{9.4}$$

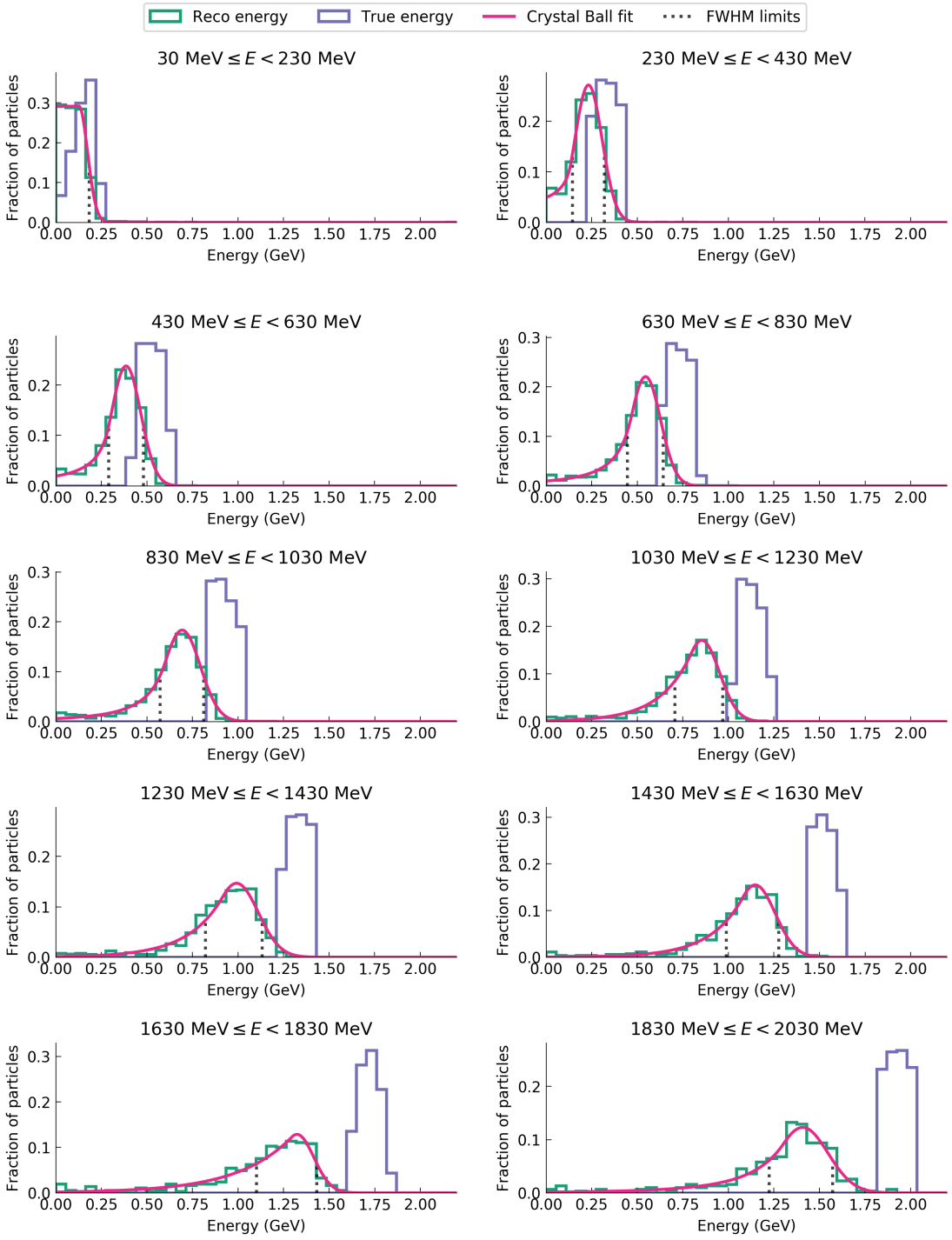


Figure 9.4: True and reconstructed shower energy in 10 equally spaced true energy bins for showers from the validation dataset, along with a Crystal Ball function fit. The FWHM of the fit is indicated by dotted lines.

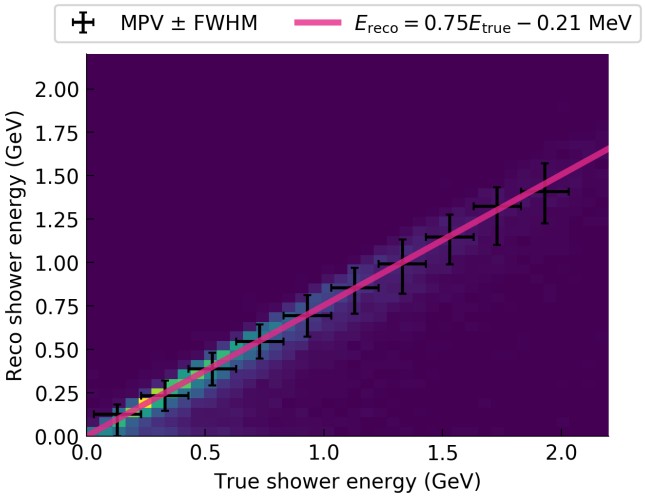


Figure 9.5: The shower energy calibration curve for showers from the validation dataset, given the most probable values of the Crystal Ball fit shown in Figure 9.4. The y error bars here come from the FWHM of the fitted distribution; the x error bars indicate the true energy bin width. The pink line shows a linear fit to the most probable values.

The distribution of E_{frac} can be characterized in each energy bin, as shown in Figure 9.6, again with a Crystal Ball fit. The typical way to proceed would be to plot the standard deviation σ from the Gaussian RHS of the Crystal Ball function. However, this fails to capture any information about the power law part of the distribution, which is where most of the uncertainty originates and indeed where most of the benefit of the algorithms evaluated in this chapter are found. We therefore propose a nonparametric alternative: we take the median absolute deviation (MAD) for the distribution, which is defined as

$$\text{MAD} := \text{median}\left(|X_i - \text{median}(X)|\right) \tag{9.5}$$

for some univariate data set $X_1, X_2, \cdots X_n$. The MAD of E_{frac} constitutes the energy resolution measurement and the plot of MAD versus E for some true track or shower energy E is used to compare algorithms. For a normal distribution with standard deviation σ, $\text{MAD} = \Phi^{-1}(3/4)\sigma \approx 0.674\sigma$, where Φ^{-1} is the quantile function for the standard normal distribution [165].

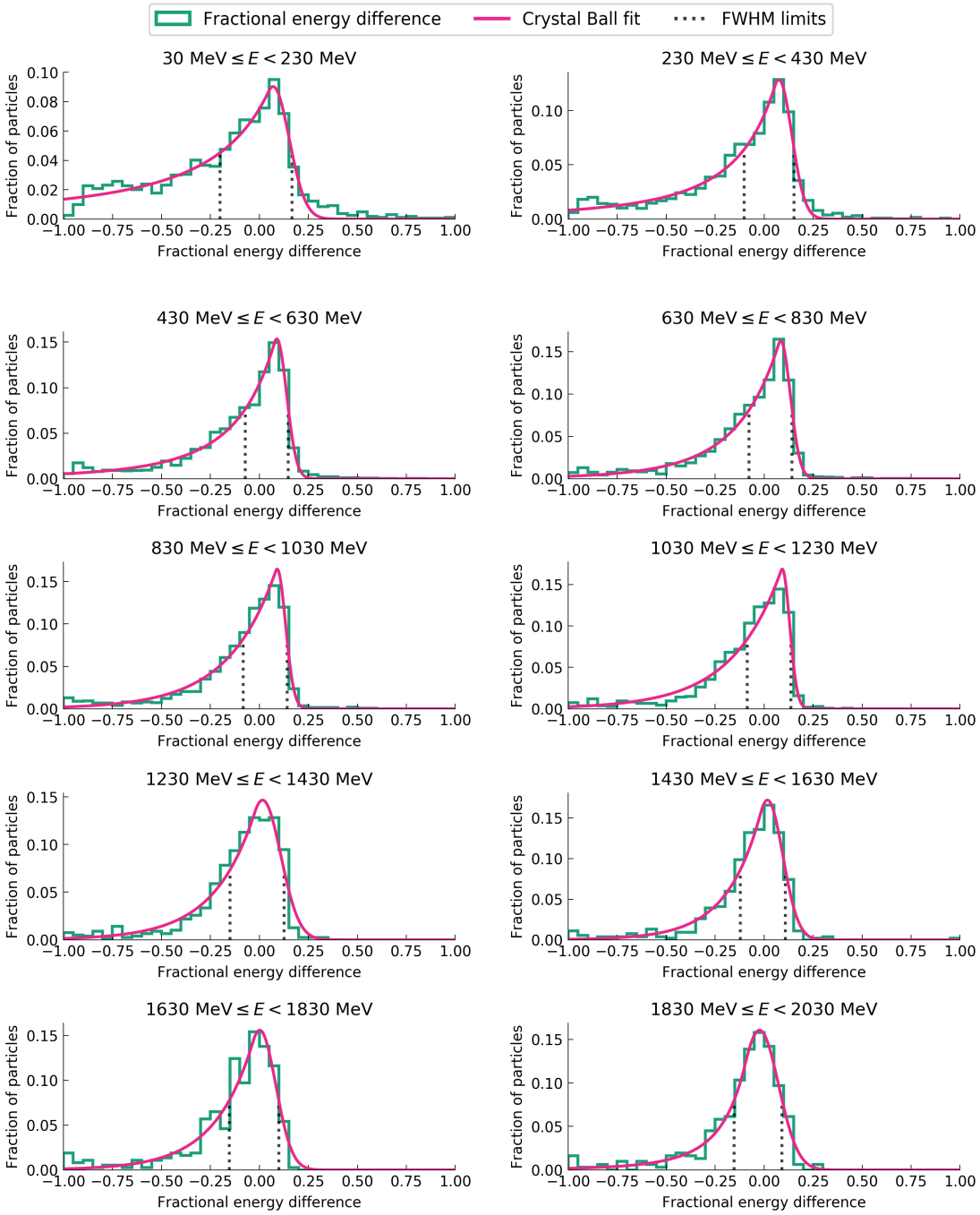


Figure 9.6: The fractional energy differences for showers from the validation dataset following calibration in each of the 10 true energy bins, along with a Crystal Ball fit. The FWHM is indicated by dotted lines.

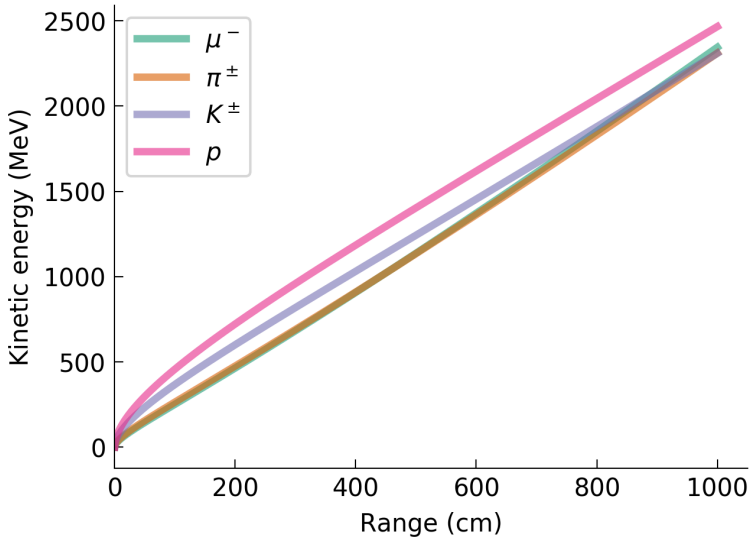


Figure 9.7: Theoretical calculation of kinetic energy from range in a MicroBooNE-like LArTPC for muons, charged pions, charged kaons and protons. This is based on the theory developed in Chapter 6, namely numerical integration of Equation 6.4 using our package `bethe-faster` [132].

9.3 Estimating track energy

The procedure used in this section follows that in [112]. Properties of the MicroBooNE LArTPC can be used to generate an estimate of a track-like particle's energy given its range. This is done via numerical integration of the Bethe-Bloch equation (see Equation 6.2) incorporating relevant detector effects.[2] In practice, we generate lookup tables using our `bethe-faster` [132] to obviate the need for repeated numerical integration, which can be computationally expensive. The resulting energies from range are shown in Figure 9.7 for a variety of track-like particles under MicroBooNE conditions.

The resulting energy is clearly a function of the particle mass, so the accuracy of PID has a direct bearing on the energy reconstruction quality. The similarity between the two particles termed MIPs (μ^- and $\pi^\pm$) is manifest. Kaons are excluded from analysis as they are relatively rare in MicroBooNE compared with other tracks. After performing the energy-from-range calculation, an analogous calibration process is performed as described for showers, this time split into 5 buckets. For each vertex/PID algorithm,

[2]For a full treatment of this, see Chapter 6.

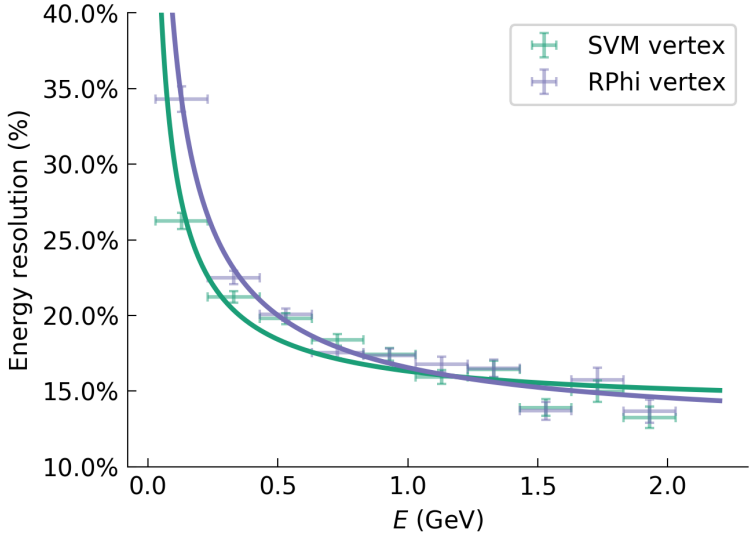


Figure 9.8: The calorimetric energy resolution plot for showers from the validation dataset using the SVM or RPhi neutrino interaction vertex selection algorithms.

all the steps of the track/shower calibration process are repeated to facilitate a fair comparison.

9.4 Results and discussion

9.4.1 Effect of neutrino interaction vertex selection algorithm

The shower energy resolution plot of MAD(E_{frac}) versus E is given in Figure 9.8 for the SVM and RPhi neutrino interaction vertex selection algorithms. For ease of comparison, we also fit an calorimetric energy resolution equation:

$$\text{MAD}(E_{\text{frac}}) = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{9.6}$$

where

- a, b and c are fit parameters representing the stochastic, noise and constant error term, respectively; and,
- $\oplus$ represents addition in quadrature.

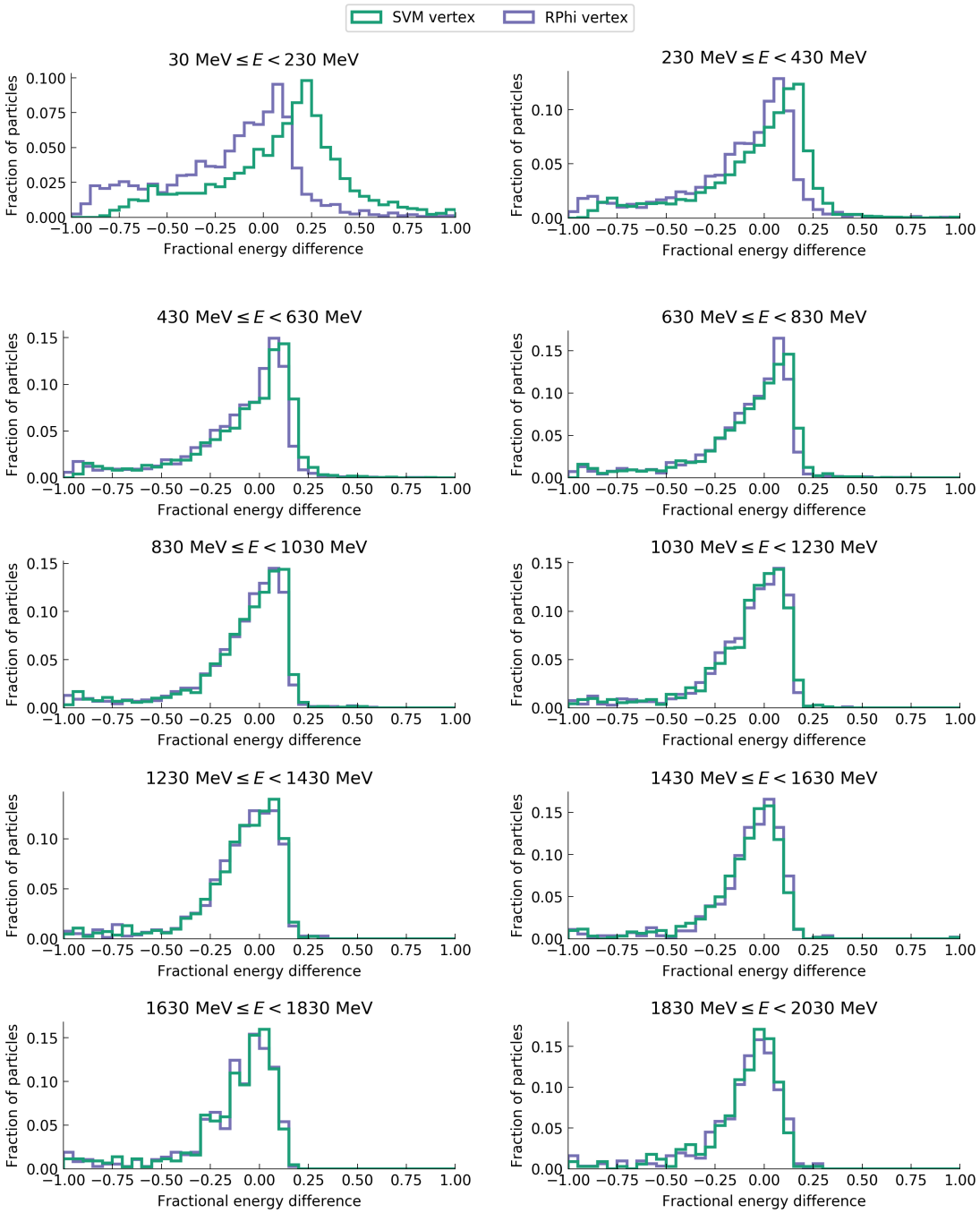


Figure 9.9: Fractional energy differences for showers from the validation dataset using the SVM or RPhi neutrino interaction vertex selection algorithms.

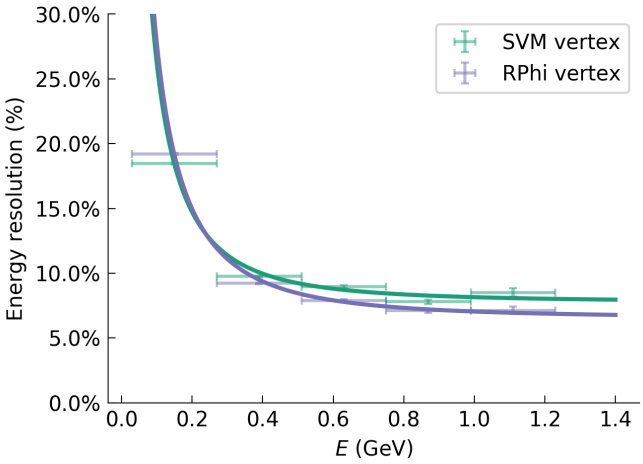


Figure 9.10: The calorimetric energy resolution plot for tracks from the validation dataset using the SVM or RPhi neutrino interaction vertex selection algorithms.

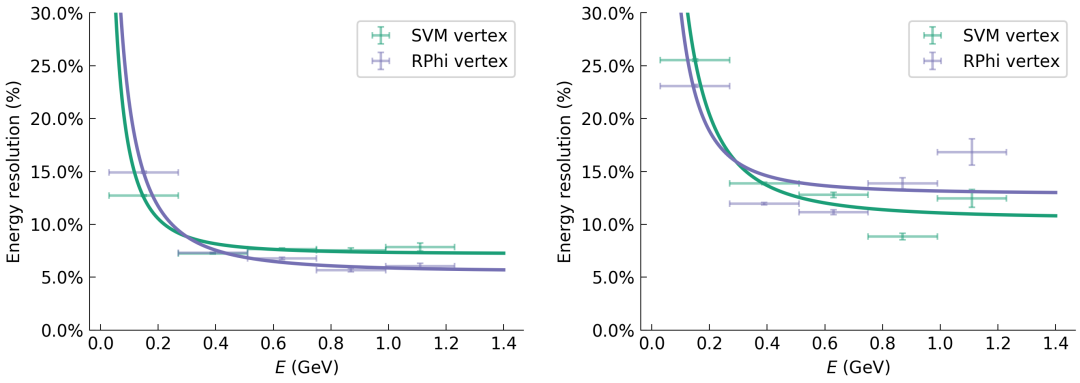


Figure 9.11: SVM/RPhi vertex algorithm comparison using the calorimetric energy resolution plot for tracks from the validation dataset. This is further split into quasi-elastic charged current BNB simulation with a single muon in the final state (`SINGLE_MU`, left) and all other channels (`EX_SINGLE_MU`, right).

The fractional energy differences in each bin are given in Figure 9.9. There appears to be significant improvement to the energy resolution for showers at lower energies when using SVM algorithm rather than RPhi, which equalizes as the energy increases to above $\sim$ 1GeV. This tallies with improvements to the CC $e^- + p$ channel seen for $E < 1$GeV in Chapter 5.

The effect of the vertex algorithm on track energy resolution is illustrated in Figure 9.10, along with the fractional energy differences in Figure 9.12. Track energy resolution

is mostly unaffected by the choice of vertex algorithm at low energies, with degradation due to the SVM algorithm apparent for $E > 0.5\,\text{GeV}$. Degradation at high energies can again be expected given the results in Chapter 5, though the improvement to low-energy track reconstruction does not appear to have a significant effect on the energy resolution. This is potentially due to the dominance of other sources of error, such as poor calorimetric reconstruction for short tracks. To probe this further, we subdivide Figure 9.10 into single muons (`SINGLE_MU`) and all other channels (`EX_SINGLE_MU`). From Chapter 5, we would expect reduced performance for `SINGLE_MU` at high energies, along with an overall improvement for `EX_SINGLE_MU` channels. This expectation is confirmed by Figure 9.11. For single muons, the likely source of performance degradation at high energies is the presence of δ-rays in the middle of tracks (see Chapter 5).

9.4.2 Effect of PID algorithm

The effect of the PID algorithm choice on track energy resolution is illustrated in Figure 9.13, along with the fractional energy differences in Figure 9.14.

Consistent with results in Chapters 7 and 8, improvements to PID quality across the energy spectrum translate into finer energy resolution. PidNet consistently outperforms PIDA and FOMA, with FOMA performing slightly better than PIDA. The relatively small scale of the changes may reflect the dominance of other modes of calorimetric error that contribute to the smearing of the distribution, such as split/merged tracks, poor 3D length reconstruction and hit-finding issues.

9.5 Conclusions

The impact of improved neutrino interaction vertex reconstruction and particle identification on particle energy estimation was one of the motivations for the work presented in this thesis, along with the intrinsic benefits of high-quality neutrino vertices and more accurate particle tagging. As expected from Chapter 5 vertex results, we observed an improvement to energy resolution at low energies for shower and tracks excluding single muons. For single muons, we observe a degradation in performance, tallying with the decreased vertex quality for this topology, which constituted part of a performance

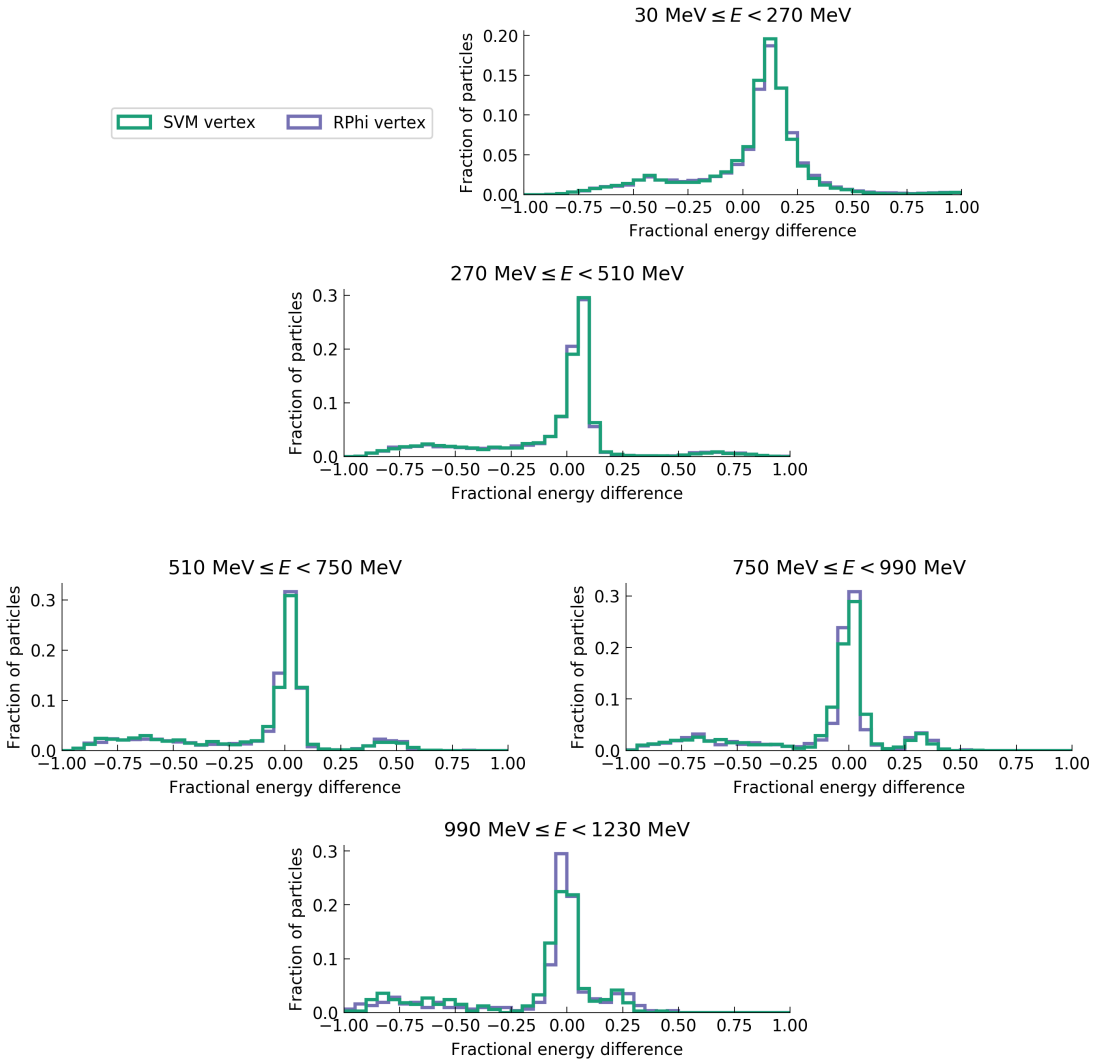


Figure 9.12: Fractional energy differences for tracks from the validation dataset using the SVM or RPhi neutrino interaction vertex selection algorithms.

tradeoff with more complex topologies. The significant increase in PID quality offered by PidNet in Chapter 8, however, translated into only modest gains in energy resolution, likely due to other reconstruction errors dominating the resolution width.

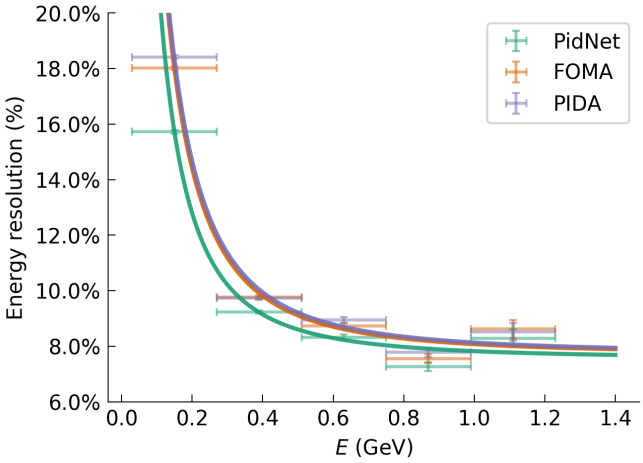


Figure 9.13: The calorimetric energy resolution plot for tracks from the validation dataset using the PID, FOMA or PidNet PID algorithms.

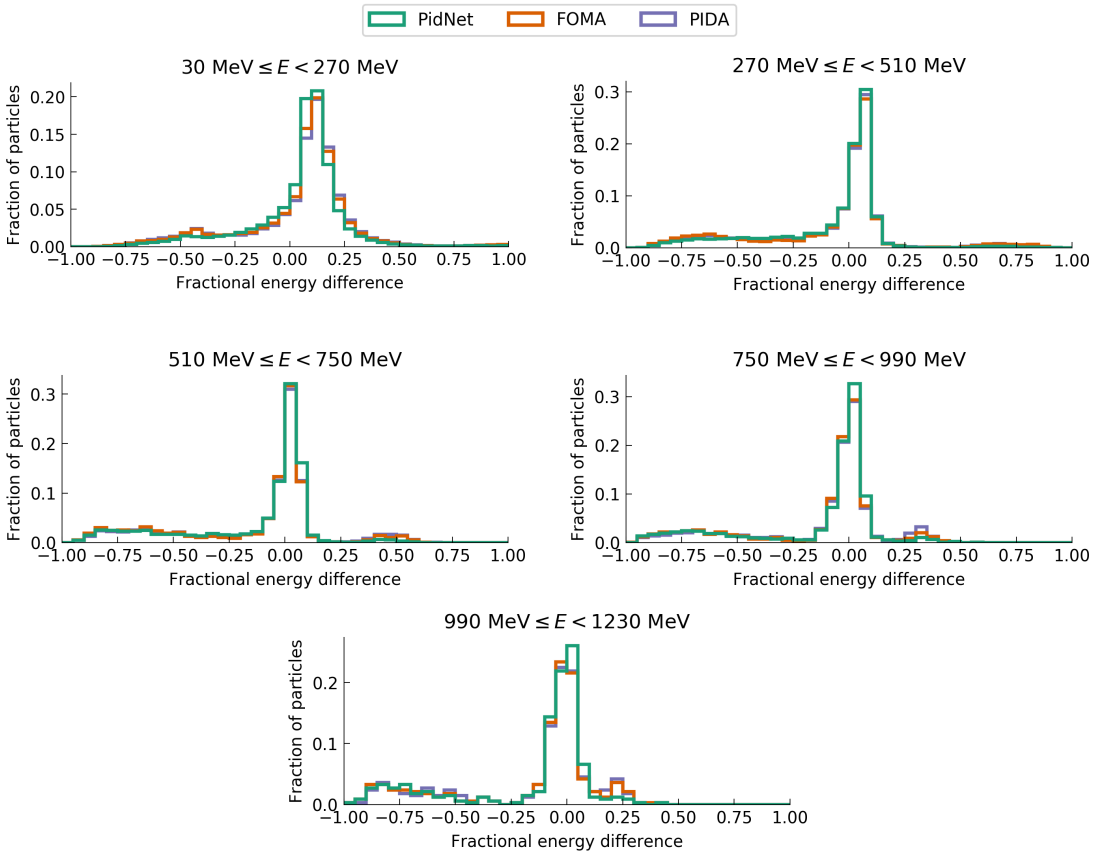


Figure 9.14: Fractional energy differences for tracks from the validation dataset using the PID, FOMA or PidNet PID algorithms.

Chapter 10

Conclusions

The primary goal of the MicroBooNE experiment is to resolve the low-energy excess of electron-like events observed at the LSND and MiniBooNE experiment, an effort which remains ongoing. Achieving this goal will represent the culmination of a vast body of research spanning the design and construction of LArTPCs, signal processing, event reconstruction and the development of novel techniques for particle identification, electron/photon discrimination and cosmic background rejection, to name a few. Beyond this flagship analysis, the same research feeds into a number of other studies, such as providing more accurate estimates of interaction cross-sections and designing algorithms for reconstructing events in LArTPCs.

We developed a new machine-learned ranking algorithm for selecting the most plausible neutrino interaction vertex candidate, the first instance of a machine learning (ML) approach used in Pandora pattern recognition. The performance benefit, particularly for more complex topologies, was demonstrated to be significant. For simulated BNB ν_μ events, the number of events with the neutrino interaction vertex reconstructed within 5 cm of the true vertex increased from 59.2% to 63.9% compared with the previous approach. For simulated intrinsic BNB ν_e events it increased from 53.6% to 59.3%. For CC resonant $\mu + \pi^0$ events, it increased from 56.3% to 71.2%; for CC resonant $e + \pi^0$ events, it increased from 45.1% to 60.9%. We noted, however, a slight decrease in vertex quality for single muons, single electrons and high-energy $\mu + Np$ events due to δ-rays and track-like/shower-like electron transitions, coupled with less reliance on simple beam

deweighting. There is scope for such issues to be corrected for at the analysis level.

Applying the theory of charged particle energy losses in liquid argon, we developed novel first- and second-order analytic low-energy approximations to modal $\mathrm{d}Q/\mathrm{d}x$ and studied their error in a LArTPC. We established that modal, rather than mean, energy losses are the most statistically justifiable in this setting and used the low-energy approximations to develop a robust track PID method called the first-order modal approximation (FOMA), employing the technique of repeated median regression. Despite requiring no tuning and being agnostic to detector parameters, the two variables produced by the model are demonstrated to have modestly superior MIP/proton discriminative power than comparable method PIDA. Without MC quality cuts, FOMA achieved a classification accuracy of 88.8% against a baseline of 66.0%, compared with 87.1% against a baseline of 65.8% for PIDA on simulated BNB ν_μ events. We performed a study on real data and observe reasonable FOMA data/MC agreement, comparable to that of PIDA.

To address the decrease in performance observed in simulated events without MC cuts using the classical PIDA/FOMA approaches, we developed a data-driven method for track PID called PidNet using deep learning. Inspired by the PIDA/FOMA data/MC domain adaptation philosophy, we drew on developments in representation learning by training a two-channel 1D convolutional autoencoder to learn 16-dimensional representations of the input calorimetric signal. We demonstrated that the method performs significantly better than both PIDA and FOMA on MIP/proton classification across a broad energy spectrum, achieving an accuracy of 93.4% against a baseline of 65.5%. We carried out a study comparing real data with simulation which showed encouraging agreement, confirming that the representations had successfully abstracted generalizable information from the calorimetric profiles.

Finally, we considered the effect of the improvements to neutrino vertex reconstruction and PID on track and shower energy estimation using standard calibration and energy-from-range methods performed in MicroBooNE analyses. We demonstrated that the new vertex selection algorithm improves shower energy resolution compared with the previous algorithm, particularly at low energies. Track energy resolution, however, appears to deteriorate somewhat in the highest considered energy range $990\,\mathrm{MeV} \leq E < 1,230\,\mathrm{MeV}$ for $\mu + Np$ events but improve for other events in this energy range. This

tallies with the slight decrease in vertex quality observed in single-muon and high-energy $\mu + Np$ events, which are abundant in the final state of ν_μ interactions. We found that PidNet offers an advantage to track energy resolution over FOMA (and FOMA over PIDA) across the energy spectrum and especially so at lower energies.

Status and next steps

Since 2017, the vertex selection algorithm presented in Chapter 5 has been the default used at MicroBooNE. This has enabled novel studies of neutrino interactions with up to six visible particles in the final state, as well as contributing directly to MicroBooNE's low-energy excess analysis. This has inspired a multitude of machine-learning-based Pandora improvements, including a SVM-based track/shower identification algorithm, a neutrino slice selection algorithm and a retrained vertex selection algorithm for reconstruction at ProtoDUNE/DUNE [166]. As a fundamental part of the reconstruction process, the vertex has far-reaching consequences for both downstream reconstruction and physics analyses, including PID and energy estimation. Future vertexing work should focus on the introduction of new features to improve predictions, such as those taking into account topological information from Pandora's particle hierarchy, and extending this work using deep learning techniques that can be made robust to data-simulation disagreement and perhaps even domain shift between experiments.

The two PID algorithms presented in Chapters 7 and 8 were finalized at the end of this thesis and have yet to be incorporated into mainstream analysis at MicroBooNE. Future work should consider the introduction of stronger physically motivated inductive biases into the architecture or objective of the deep learning algorithm, and the use of techniques that encourage robustness to data-simulation disagreement, such as adversarial domain adaptation. More broadly, all machine learning algorithms in this thesis can be tuned for analyses, topologies or experiments; the ability for users to choose from a library of trained models or fine-tune them could be low-hanging fruit for improving the quality of analyses.

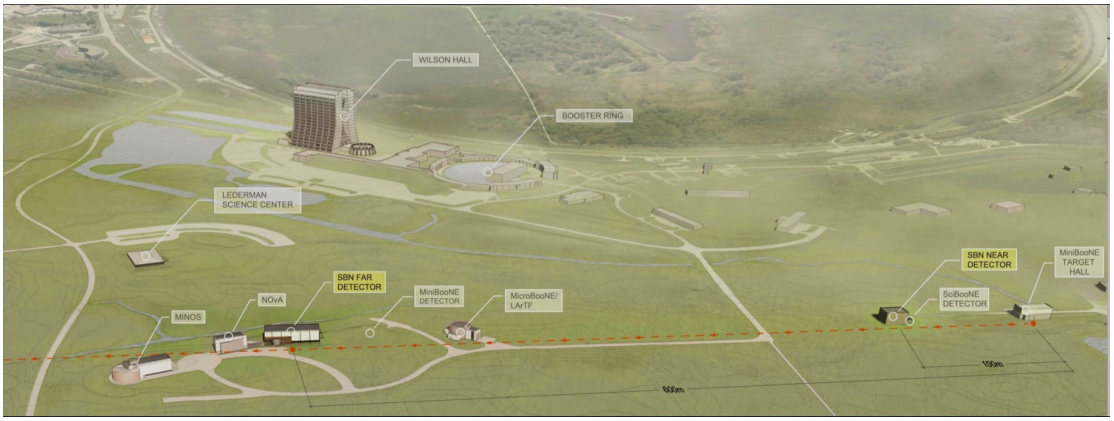


Figure 10.1: Artist's illustration of the SBN Program at Fermilab, Illinois. The BNB travels from right to left, first passing through SBND ("SBN Near Detector"), then MicroBooNE, then ICARUS T600 ("SBN Far Detector"). Figure from [20].

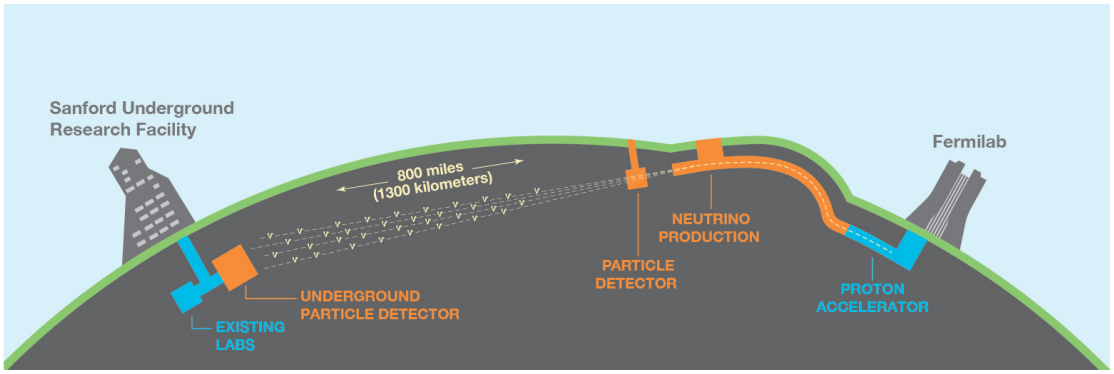


Figure 10.2: The DUNE beamline, beginning at Fermilab, Illinois, before travelling 1,300 km to a 70 kt LArTPC-based detector in the former gold mines of the Sanford Lab, South Dakota. Figure from [167].

Outlook

MicroBooNE exists as the first step in the Short-Baseline Neutrino (SBN) Program at Fermilab (see Figure 10.1) [20], preceding the Short-Baseline Far Detector (ICARUS T600) [16] and the Short-Baseline Near Detector (SBND) [17]. All three detectors employ similar LArTPC technologies and sit at different places along the BNB: SBND has an active LAr mass of 12 tons[1] and sits 100m from the BNB target; MicroBooNE 87 tons and 470m from the target; and, ICARUS T600 476 tons and 600m from the target.

[1] 1 ton = 2,000 lb here, following the US definition.

Together, these experiments offer an unprecedented opportunity to study short-baseline neutrino oscillations and constrain BNB flux measurements, interaction cross-sections and detector effects, as well as facilitating knowledge-sharing at the signal processing, reconstruction and analysis levels. The next stage in our understanding of neutrino physics will stem from the long-baseline Deep Underground Neutrino Experiment (DUNE; see Figure 10.2) [166] at Fermilab and the Hyper-Kamiokande (HK) experiment in Japan [18]. Both DUNE and HK hope to observe and measure CP violation in the leptonic sector [19], determine the ordering of the neutrino masses, observe cosmic neutrinos and search for signals of proton decay. Knowledge and techniques developed at MicroBooNE, such as event reconstruction and PID in LArTPCs, will feed directly into reconstruction and analysis at DUNE. Algorithms developed in this thesis are already being used to reconstruct real ProtoDUNE events and simulated DUNE events.

Appendix A

Further vertex selection algorithm validation

A.1 Vertex SVM performance

Figure A.1 shows a summary of the performance of the vertex-level classifier; that is, the performance of the classifier that tries to select the true vertex from vertices in the same (best) region. See Chapter 5 for full details. Figure A.2 shows the distribution of features used in the vertex selection SVM for the true best vertex and for the other vertices in the region.

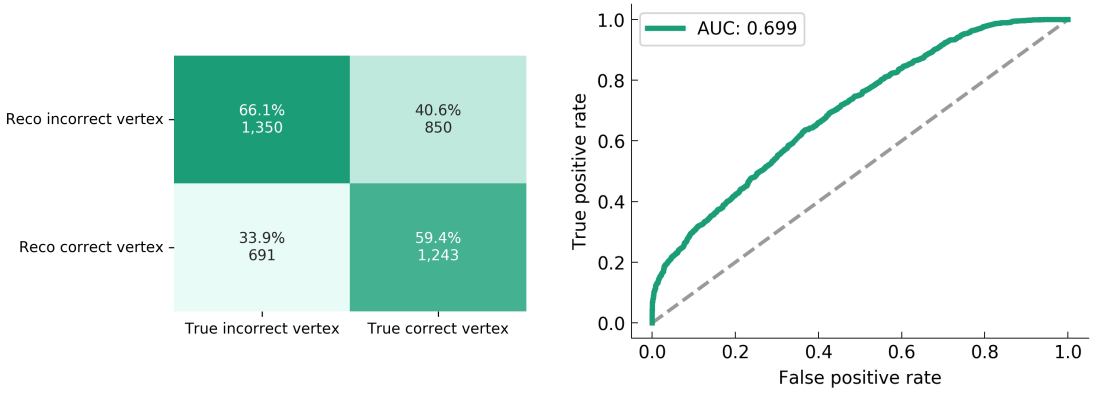


Figure A.1: The performance of the vertex classifier: the confusion matrix (left) and ROC curve (right) across the simulated BNB ν_μ spectrum.

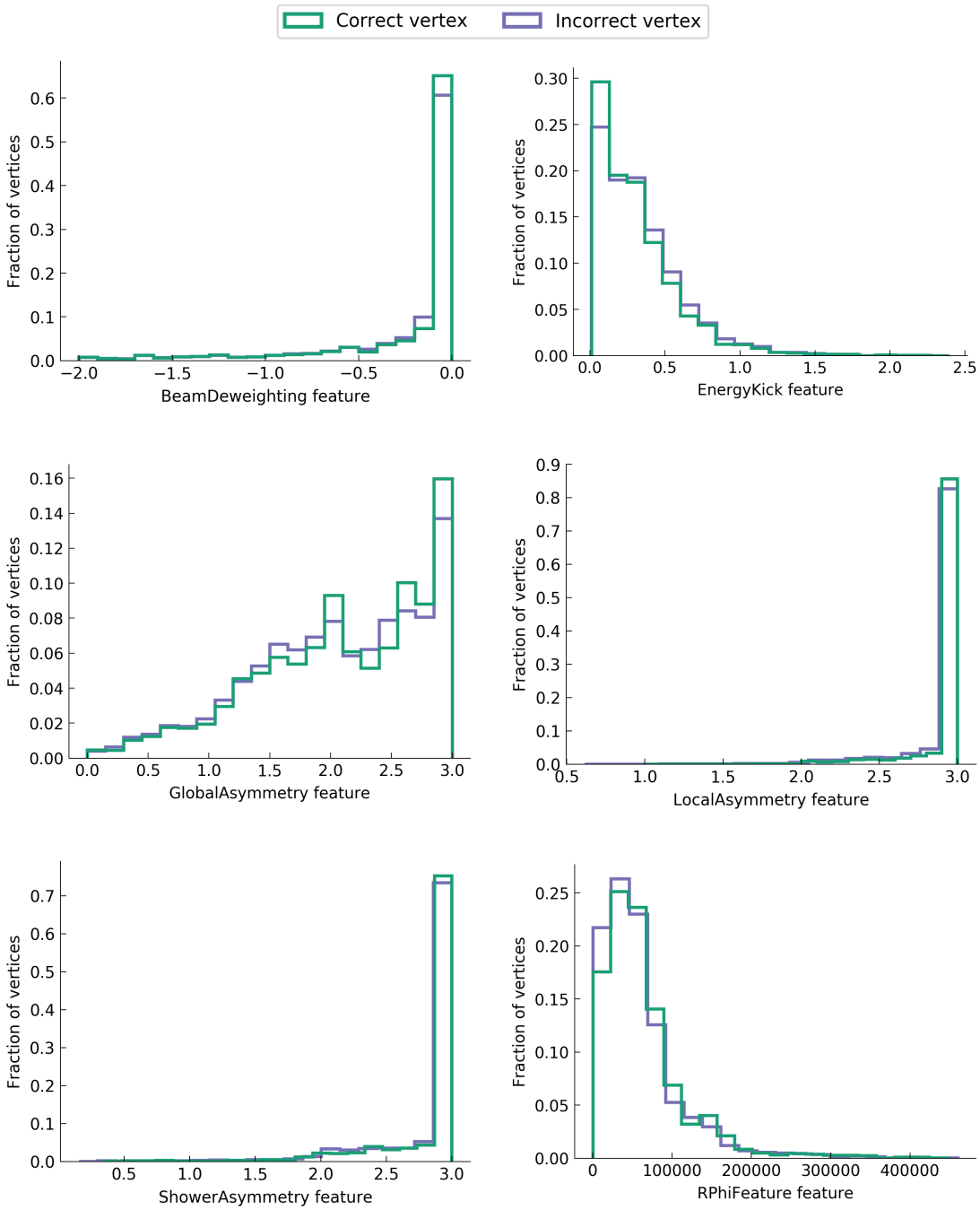


Figure A.2: The distributions of features used in the single vertex selection algorithm for the correct and incorrect vertex candidates across the simulated BNB ν_μ spectrum.

A.2 Detailed results

This section explores the performance of the vertex-level and region-level classifiers in greater detail. This includes MC matching purity and completeness for the BNB ν_μ spectrum (Figure A.3), MC matching purity and completeness for the BNB ν_e spectrum (Figure A.4), comparing the two vertex algorithms. The remaining plots in this section show channel-specific performance and feature distributions for channels of general interest at MicroBooNE (`CCQEL_MU`, `CCQEL_MU_P`, `CCRES_MU_PIZERO`, `CCQEL_E`, `CCQEL_E_P` and `CCRES_E_PIZERO`). For a detailed description of each plot, refer to the captions.

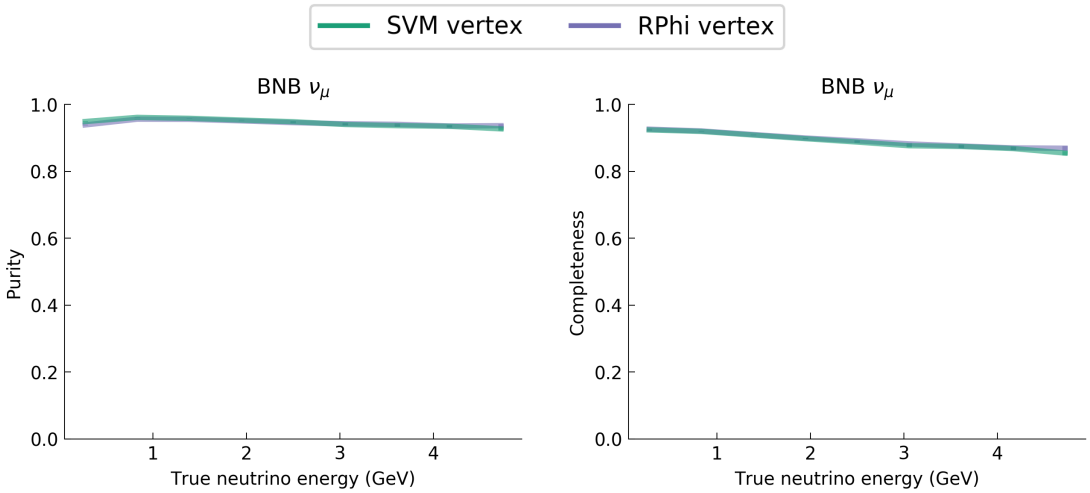


Figure A.3: The purity and completeness of the MC matches for the two vertex algorithms as a function of true neutrino energy across the simulated BNB ν_μ spectrum.

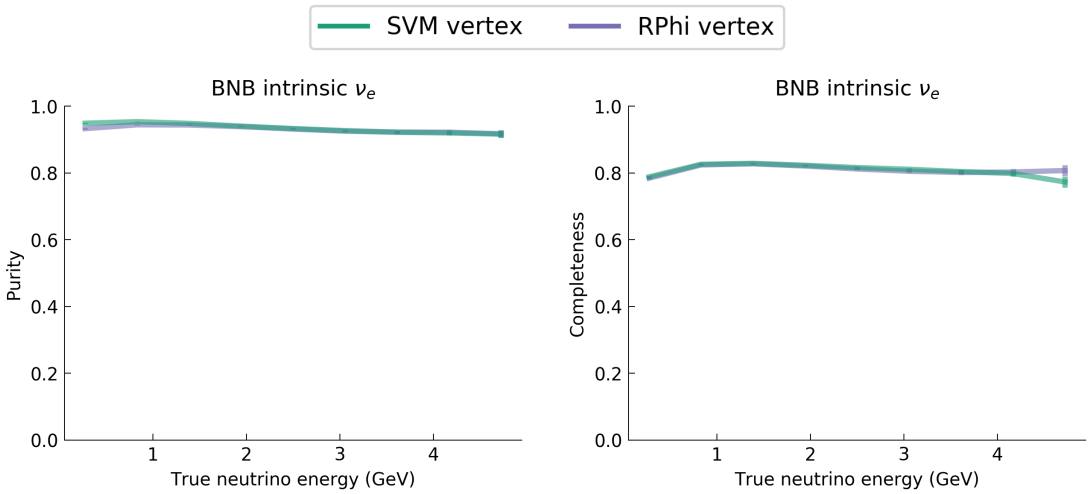


Figure A.4: The purity and completeness of the MC matches for the two vertex algorithms as a function of true neutrino energy across the simulated BNB intrinsic ν_e spectrum.

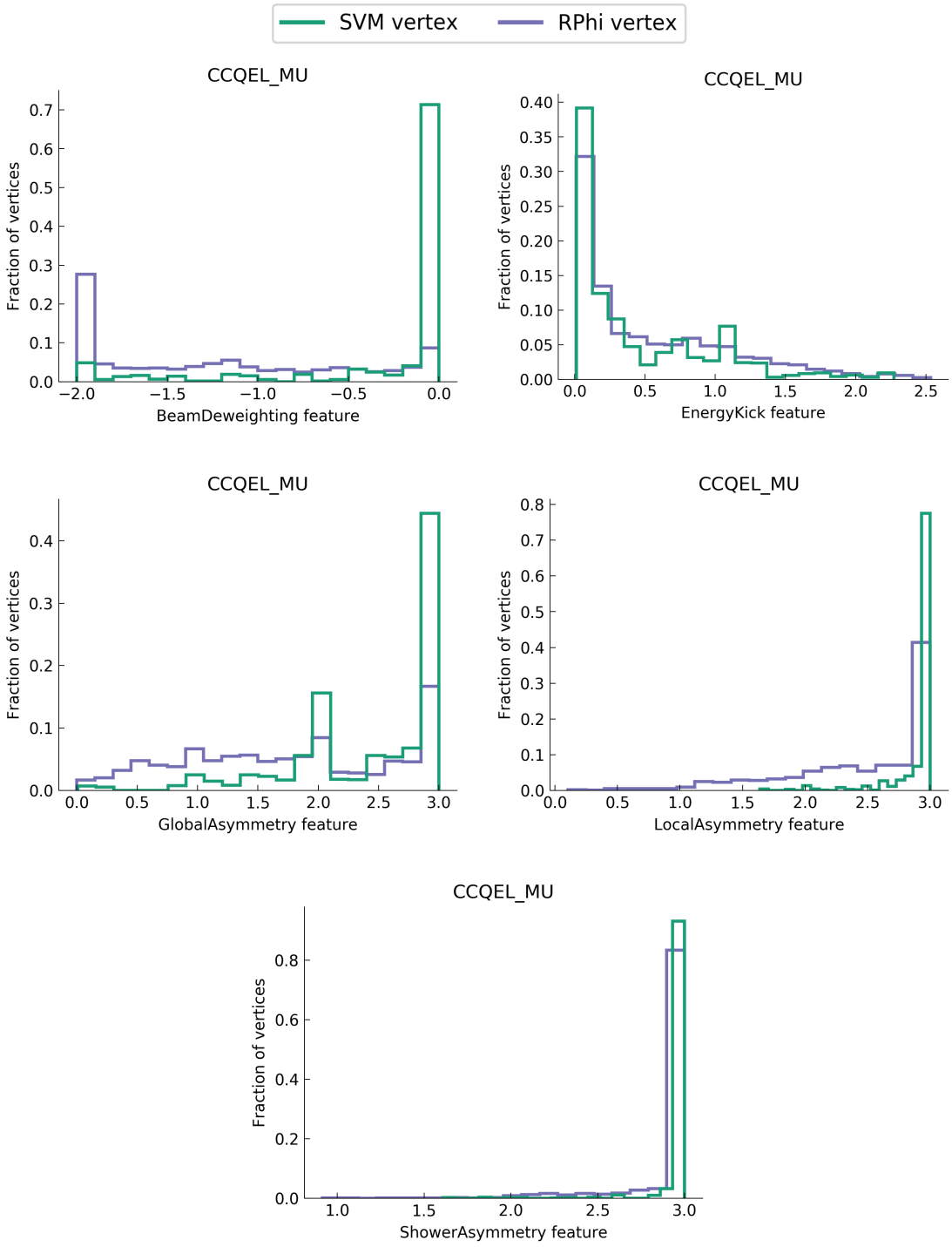


Figure A.5: The distributions of features used in the region selection algorithm for the correct and correct vertex candidates for the simulated BNB ν_μ CC quasi-elastic $\mu + 0p$ channel.

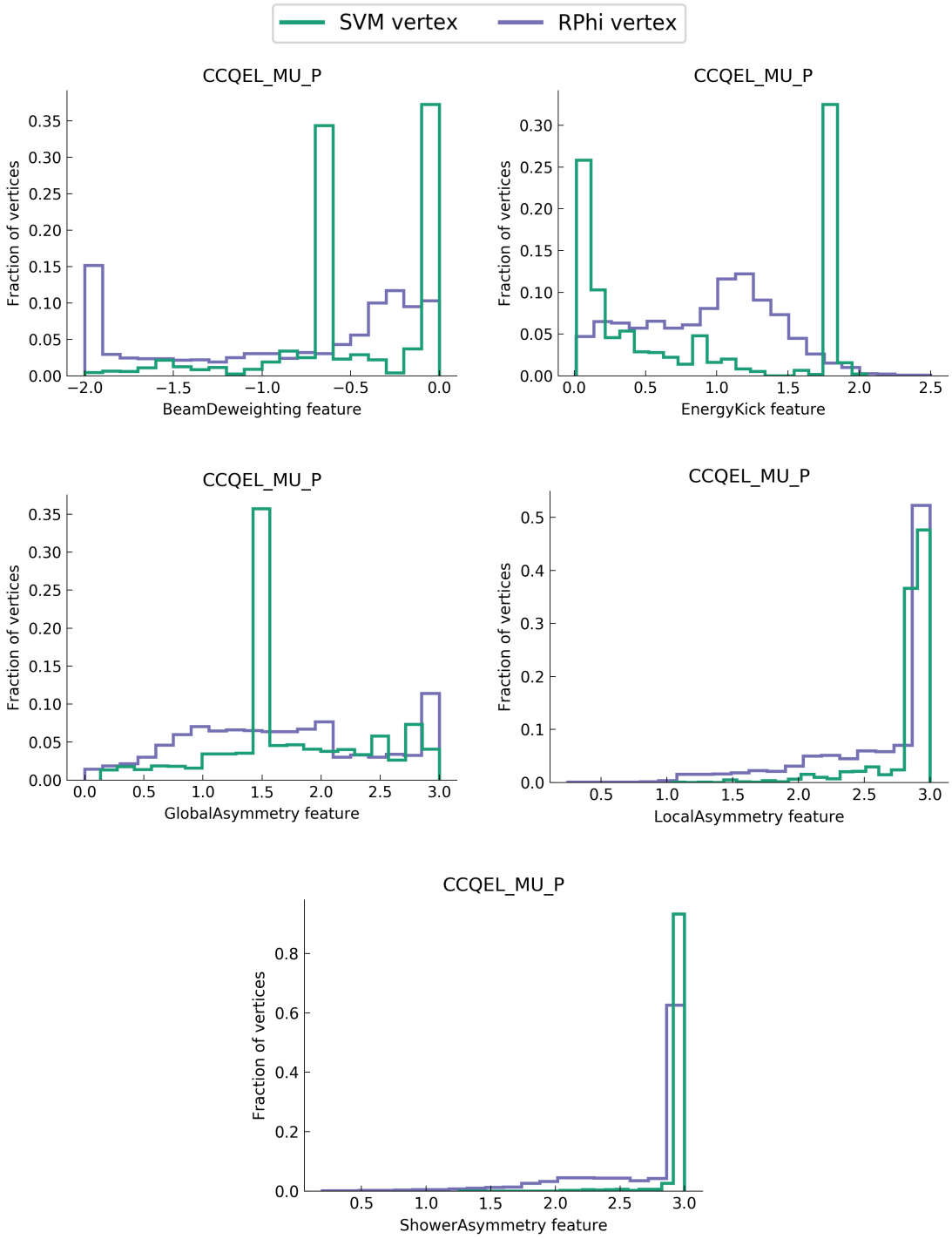


Figure A.6: The distributions of features used in the region selection algorithm for the correct and correct vertex candidates for the simulated BNB ν_μ CC quasi-elastic $\mu + 1p$ channel.

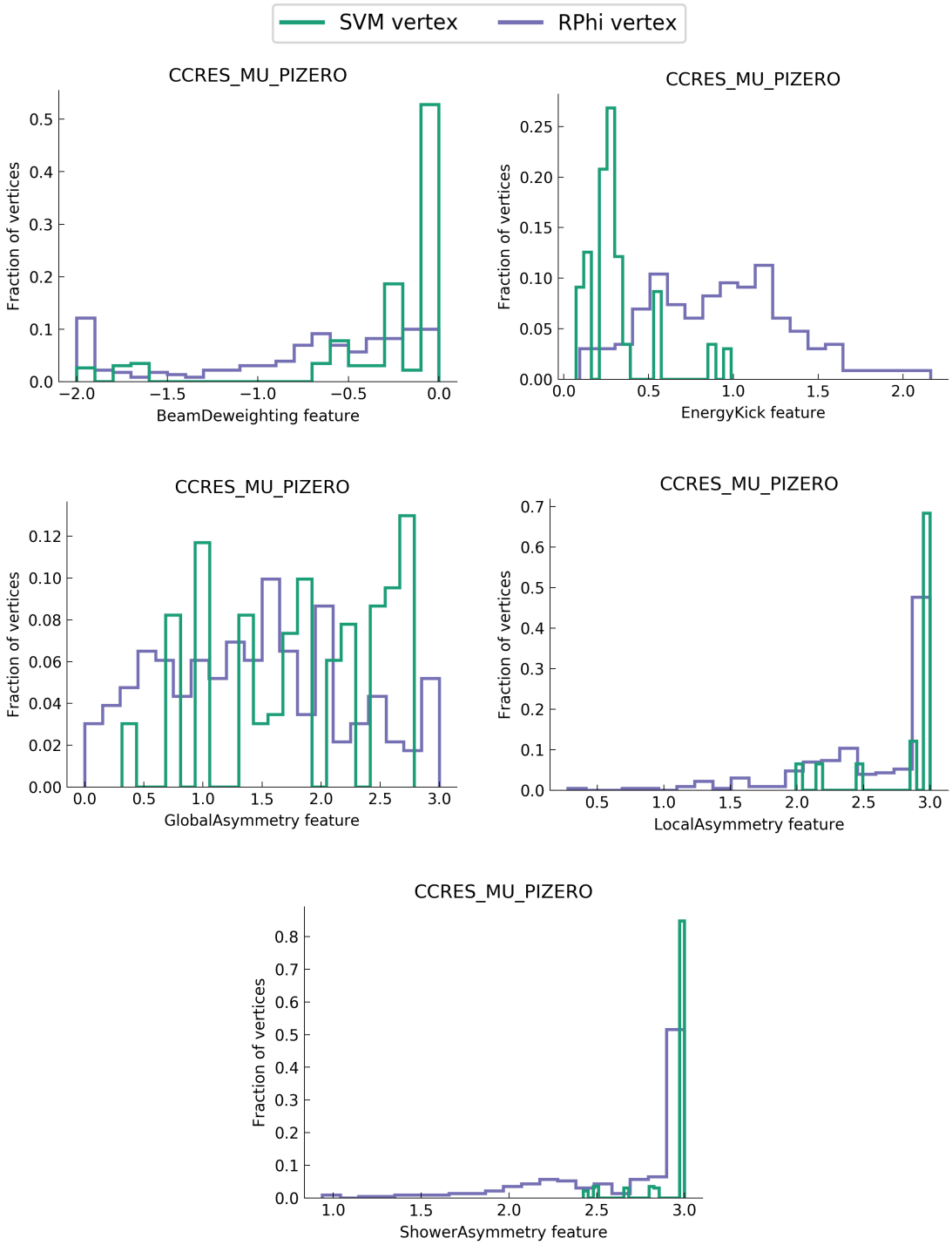


Figure A.7: The distributions of features used in the region selection algorithm for the correct and correct vertex candidates for the simulated BNB ν_μ CC resonant $\mu + \pi^0$ channel.

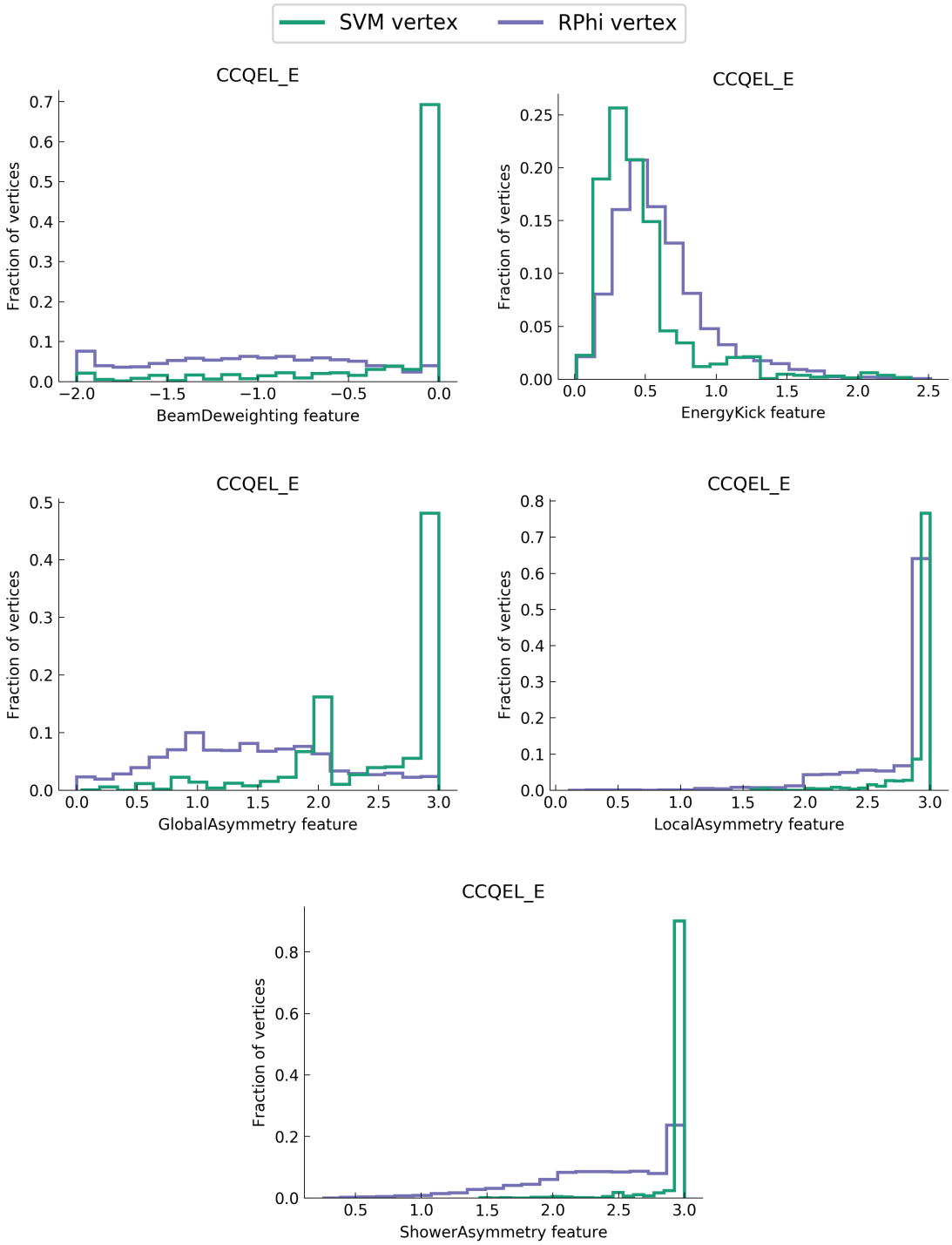


Figure A.8: The distributions of features used in the region selection algorithm for the correct and correct vertex candidates for the simulated BNB intrinsic ν_e CC quasi-elastic $e + 0p$ channel.

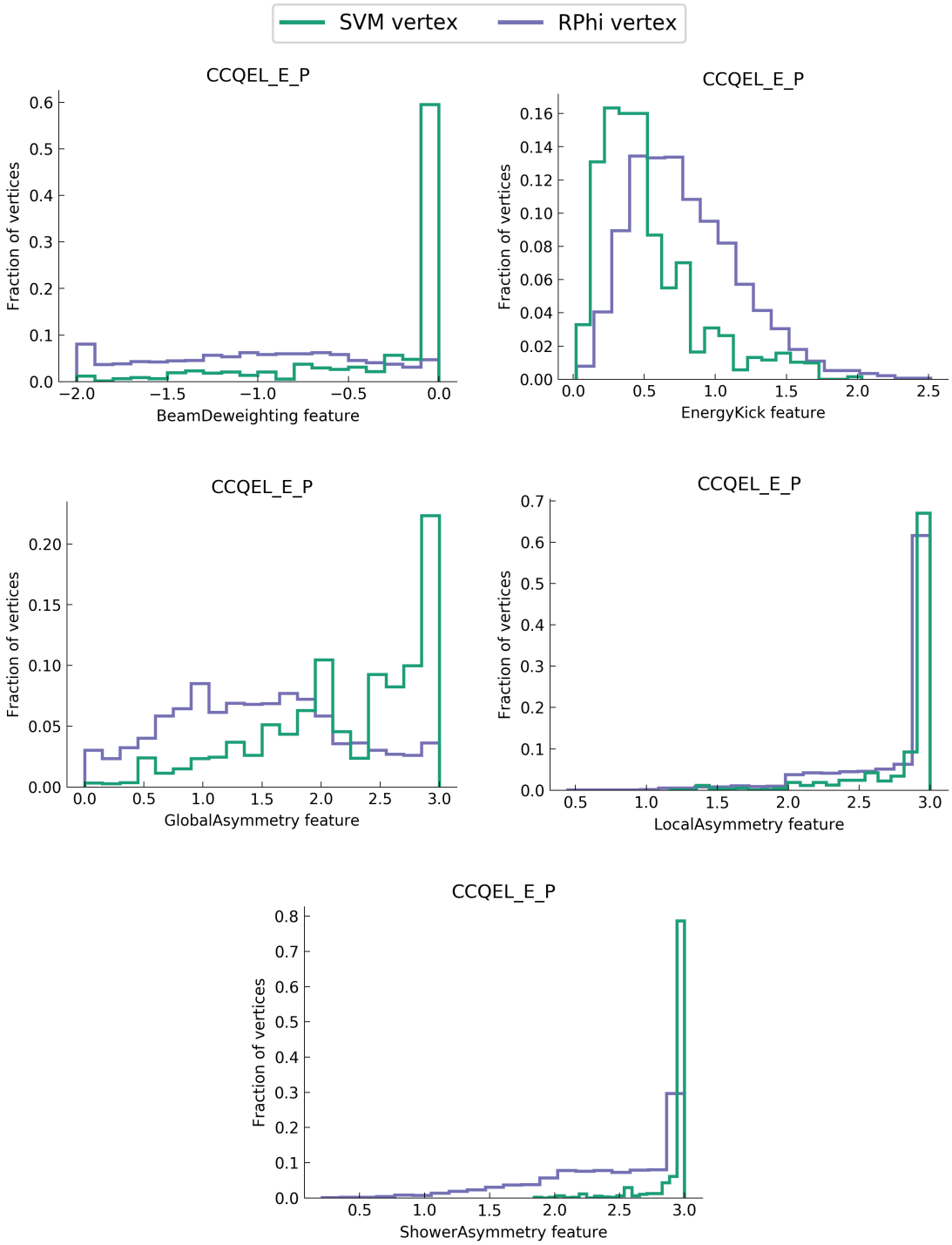


Figure A.9: The distributions of features used in the region selection algorithm for the correct and correct vertex candidates for the simulated BNB intrinsic ν_e CC quasi-elastic $e + 1p$ channel.

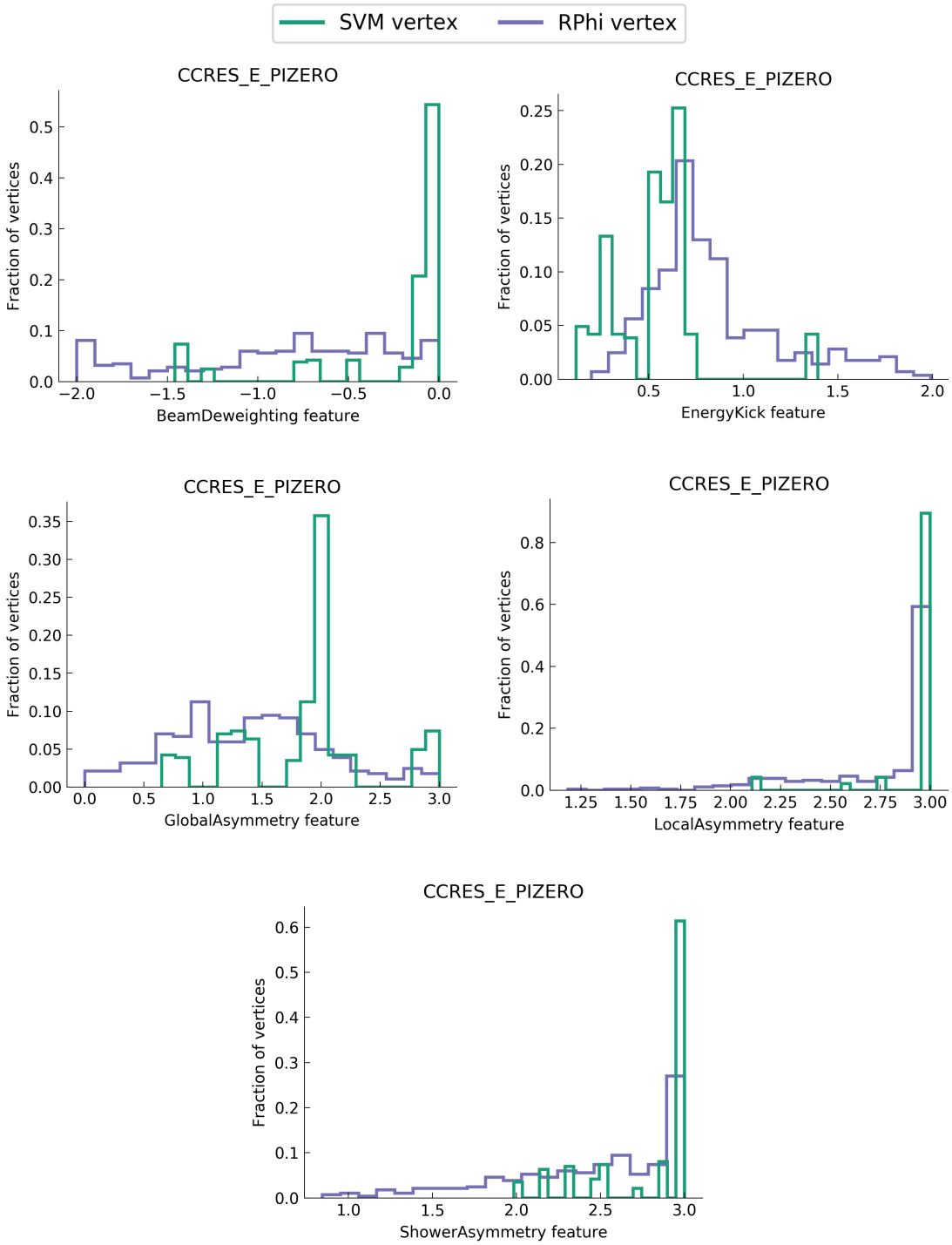


Figure A.10: The distributions of features used in the region selection algorithm for the correct and correct vertex candidates for the simulated BNB intrinsic ν_e CC resonant $e + \pi^0$ channel.

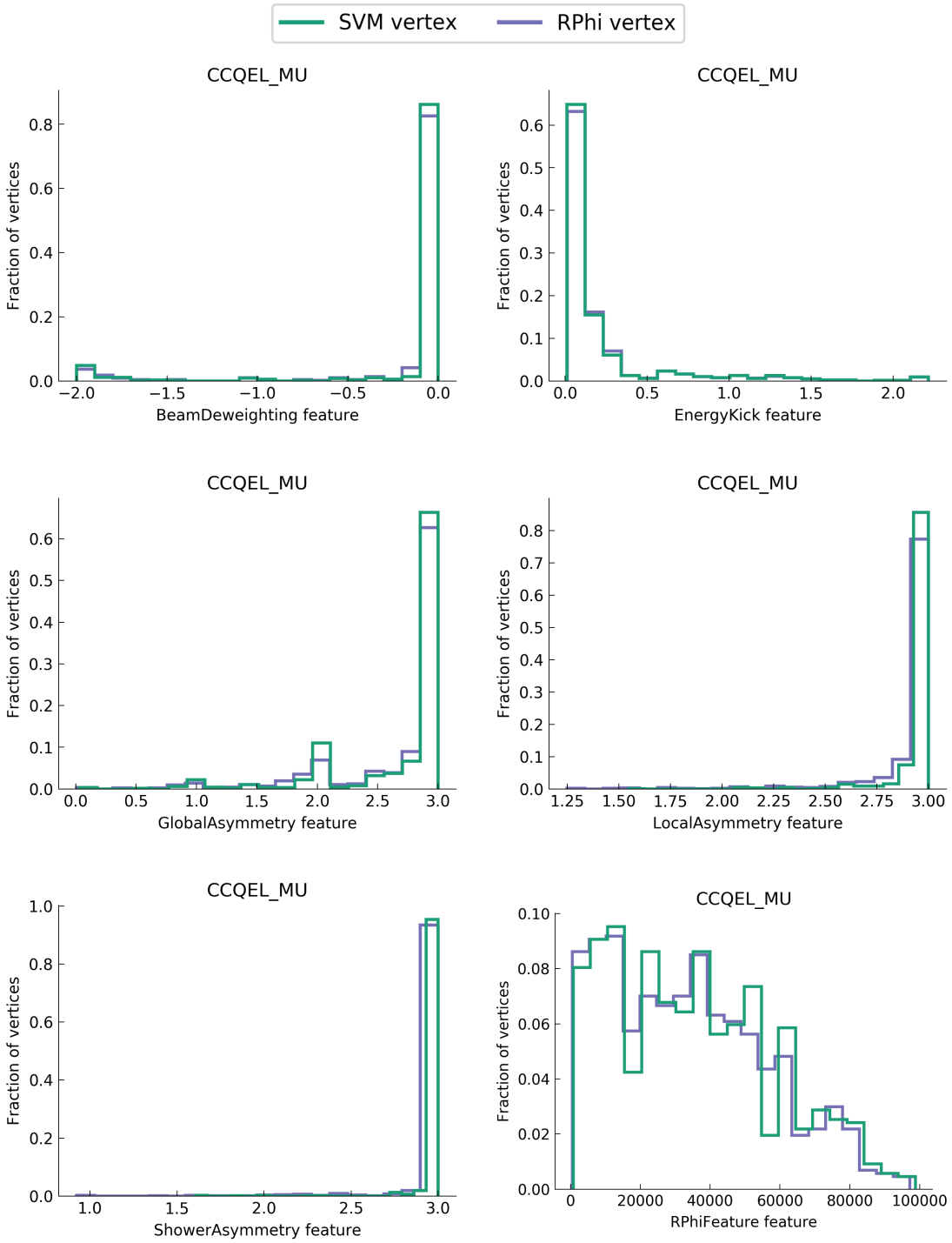


Figure A.11: The distributions of features used in the vertex selection algorithm for the correct and correct vertex candidates for the simulated BNB ν_μ CC quasi-elastic $\mu + 0p$ channel.

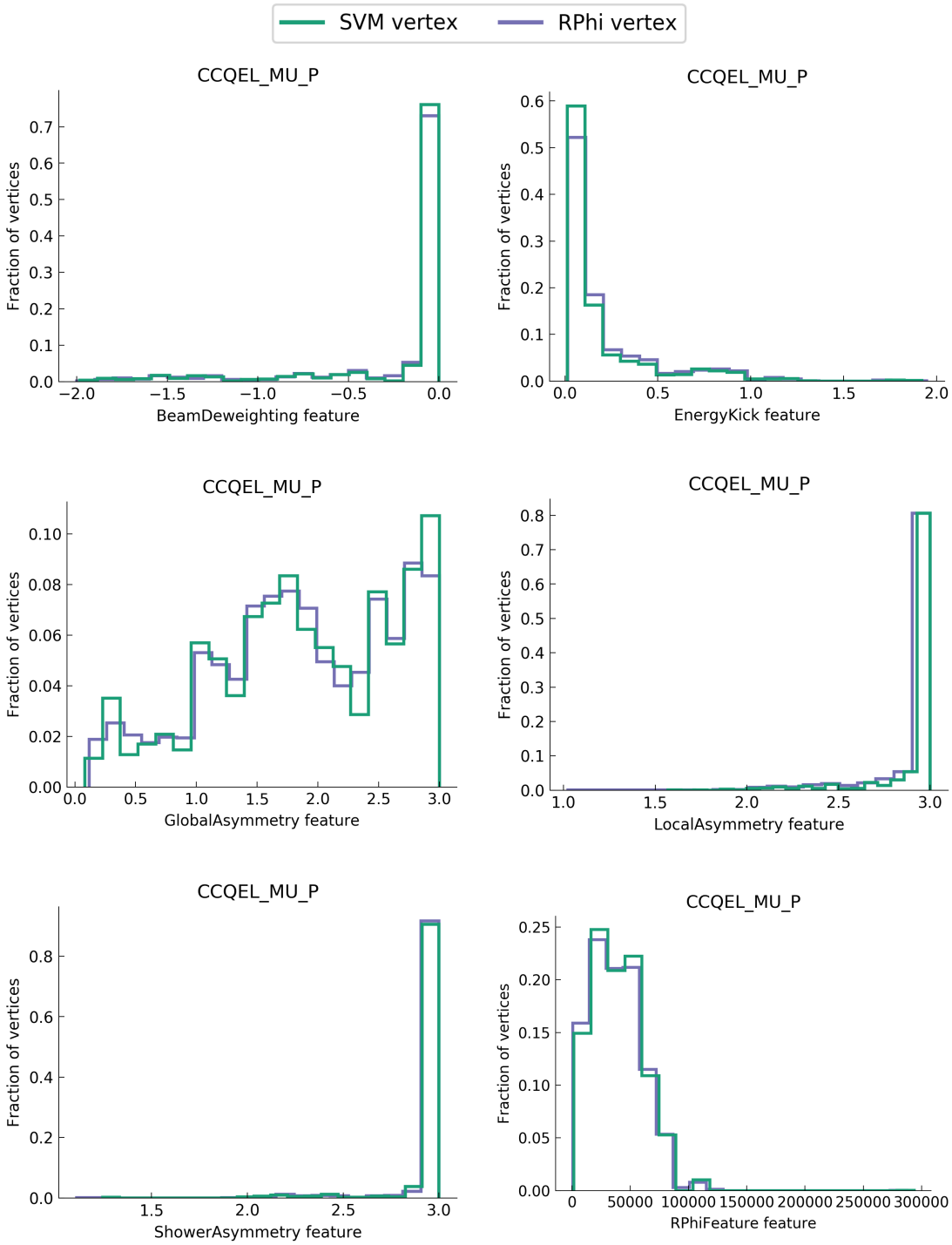


Figure A.12: The distributions of features used in the vertex selection algorithm for the correct and correct vertex candidates for the simulated BNB ν_μ CC quasi-elastic $\mu + 1p$ channel.

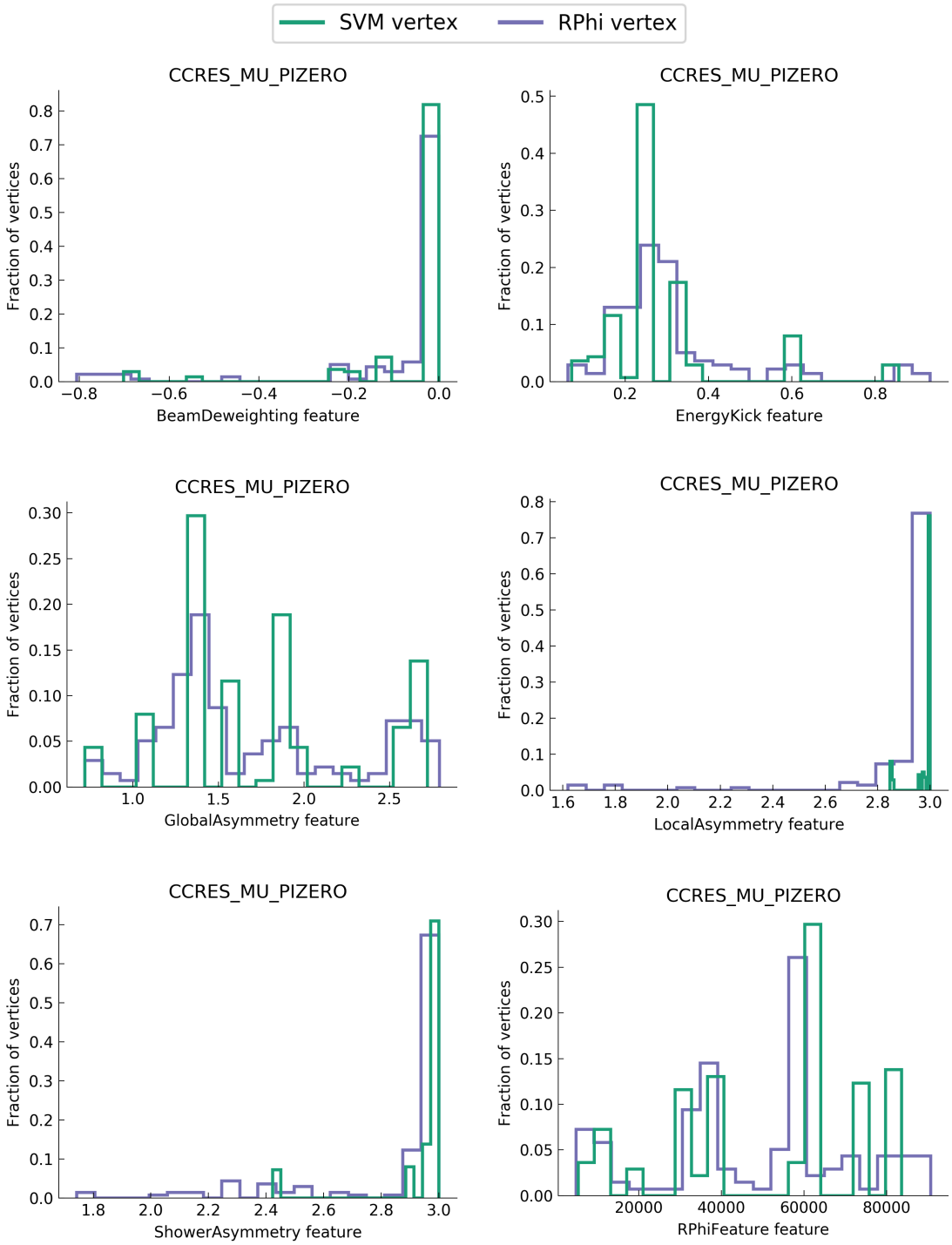


Figure A.13: The distributions of features used in the vertex selection algorithm for the correct and correct vertex candidates for the simulated BNB ν_μ CC resonant $\mu + \pi^0$ channel.

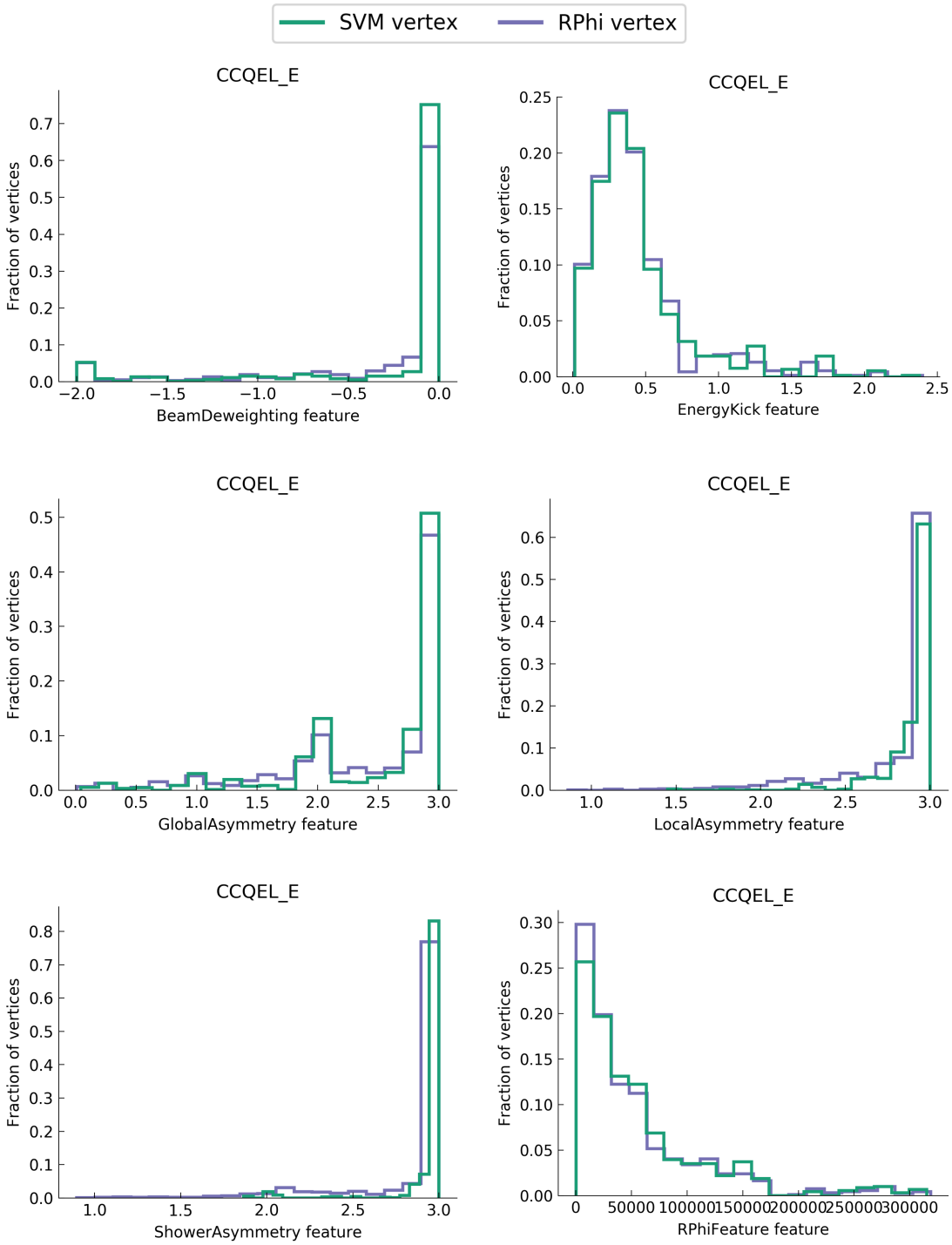


Figure A.14: The distributions of features used in the vertex selection algorithm for the correct and correct vertex candidates for the simulated BNB intrinsic ν_e CC quasi-elastic $e + 0p$ channel.

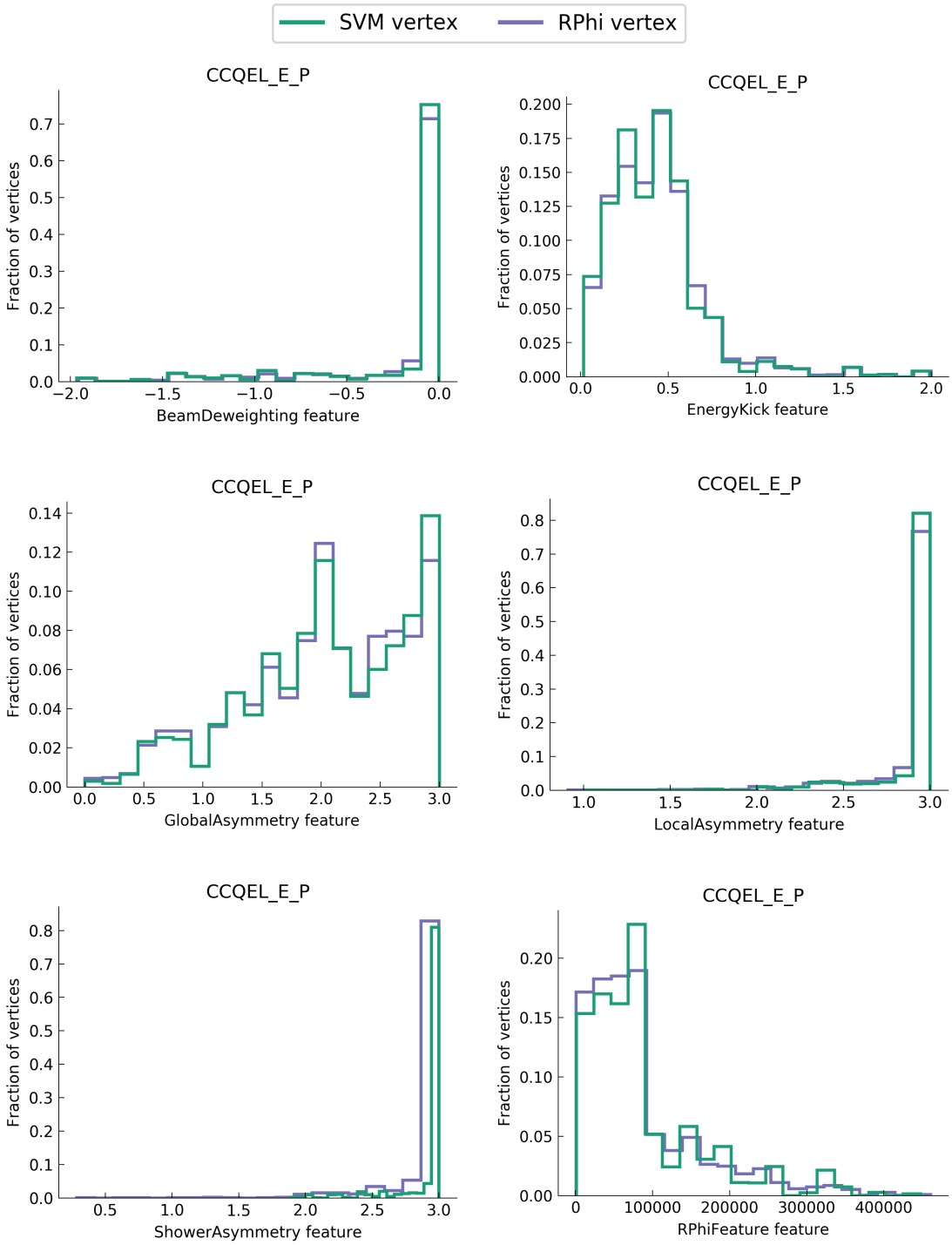


Figure A.15: The distributions of features used in the vertex selection algorithm for the correct and correct vertex candidates for the simulated BNB intrinsic ν_e CC quasi-elastic $e + 1p$ channel.

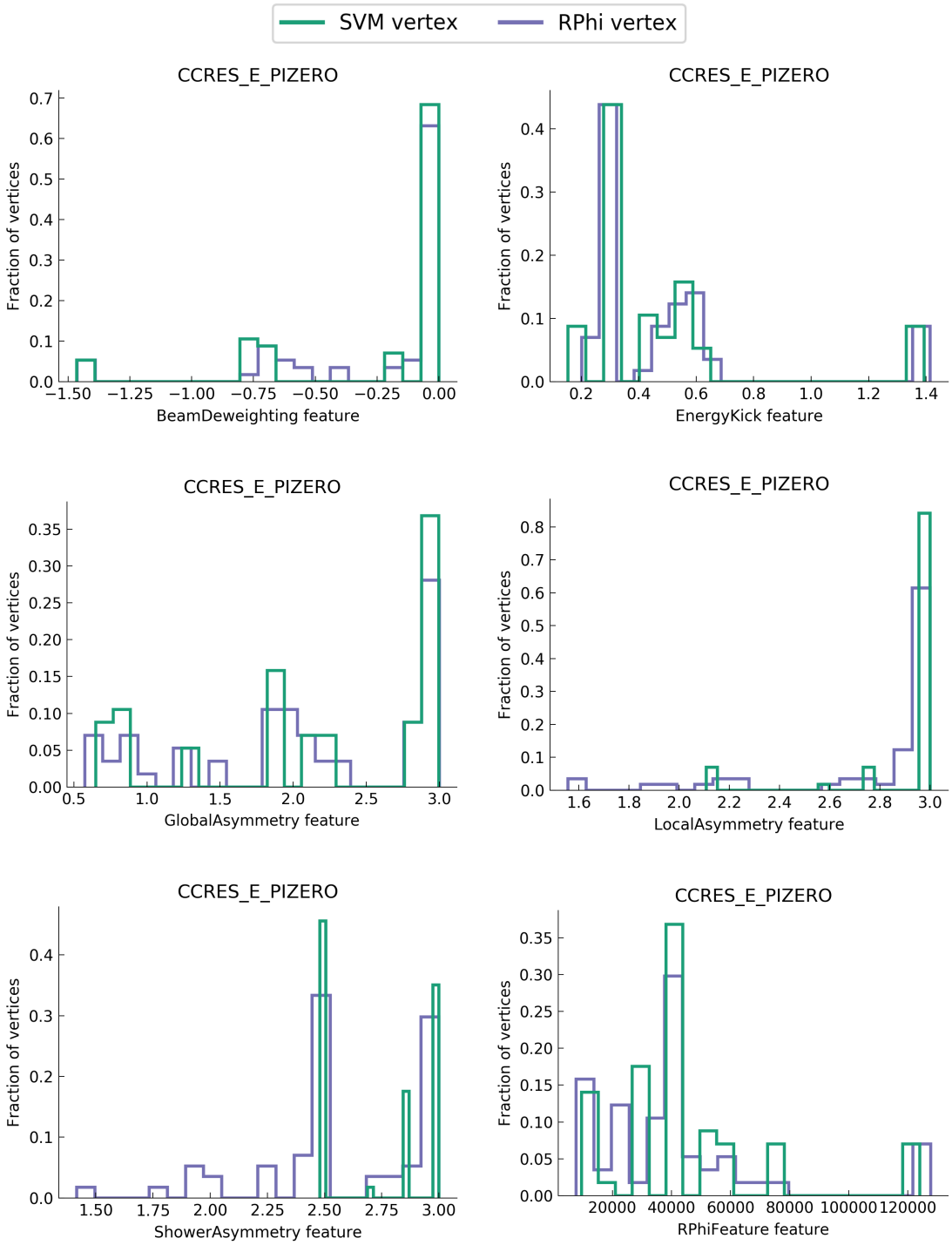


Figure A.16: The distributions of features used in the vertex selection algorithm for the correct and correct vertex candidates for the simulated BNB intrinsic ν_e CC resonant $e + \pi^0$ channel.

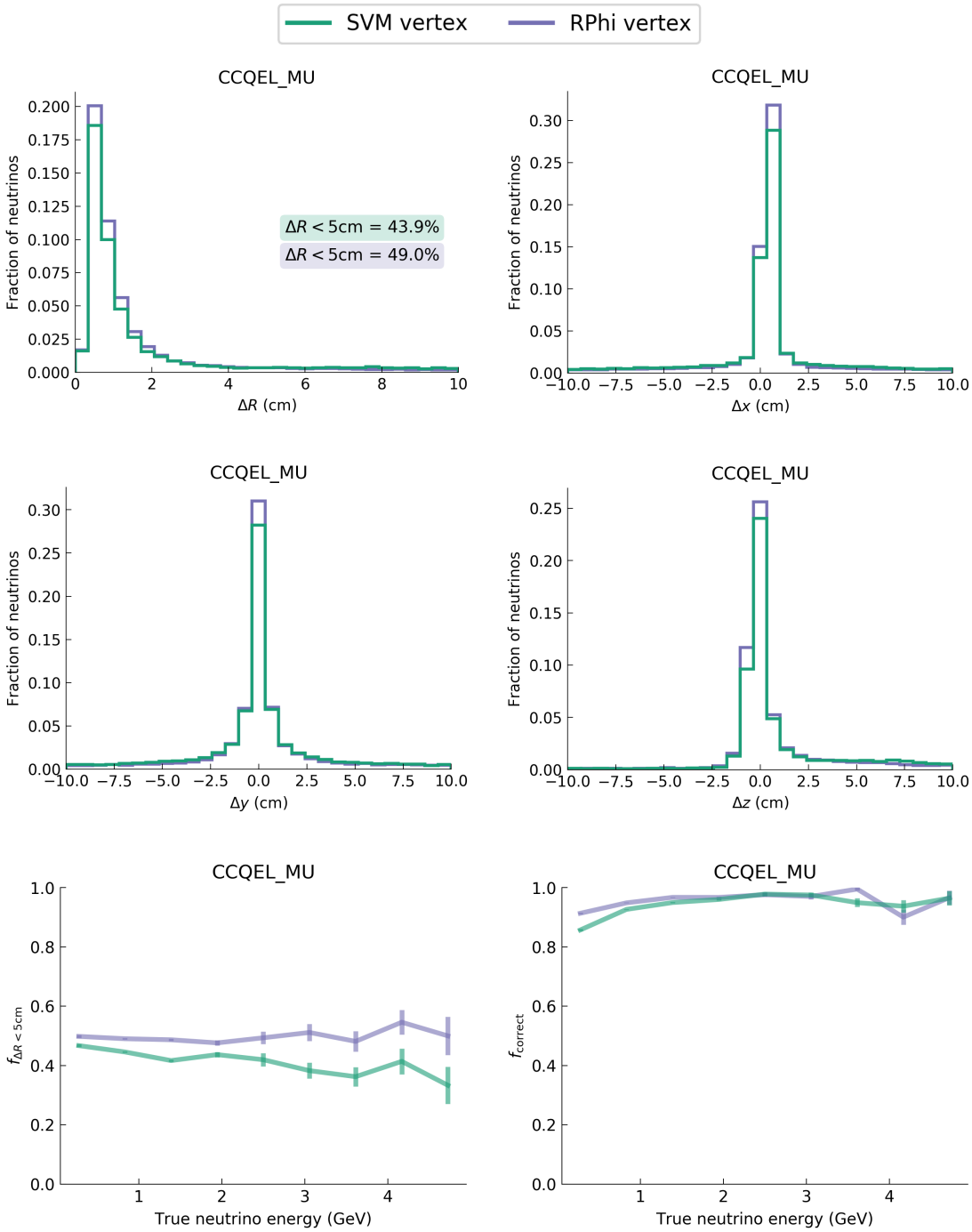


Figure A.17: Vertex reconstruction quality for the two algorithms for the BNB ν_μ CC quasi-elastic $\mu + 0p$ channel.

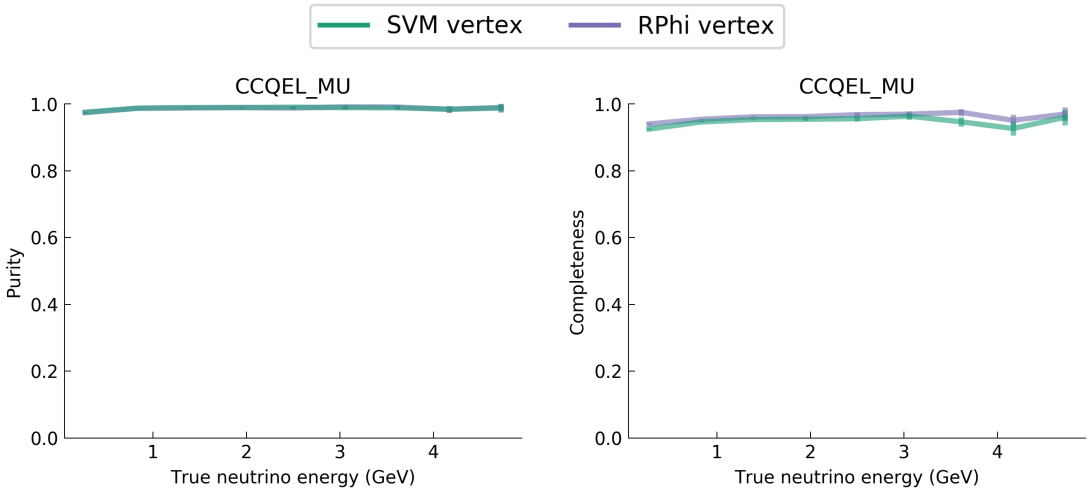


Figure A.18: MC particle matching quality for the two algorithms for the BNB ν_μ CC quasi-elastic $\mu + 0p$ channel.

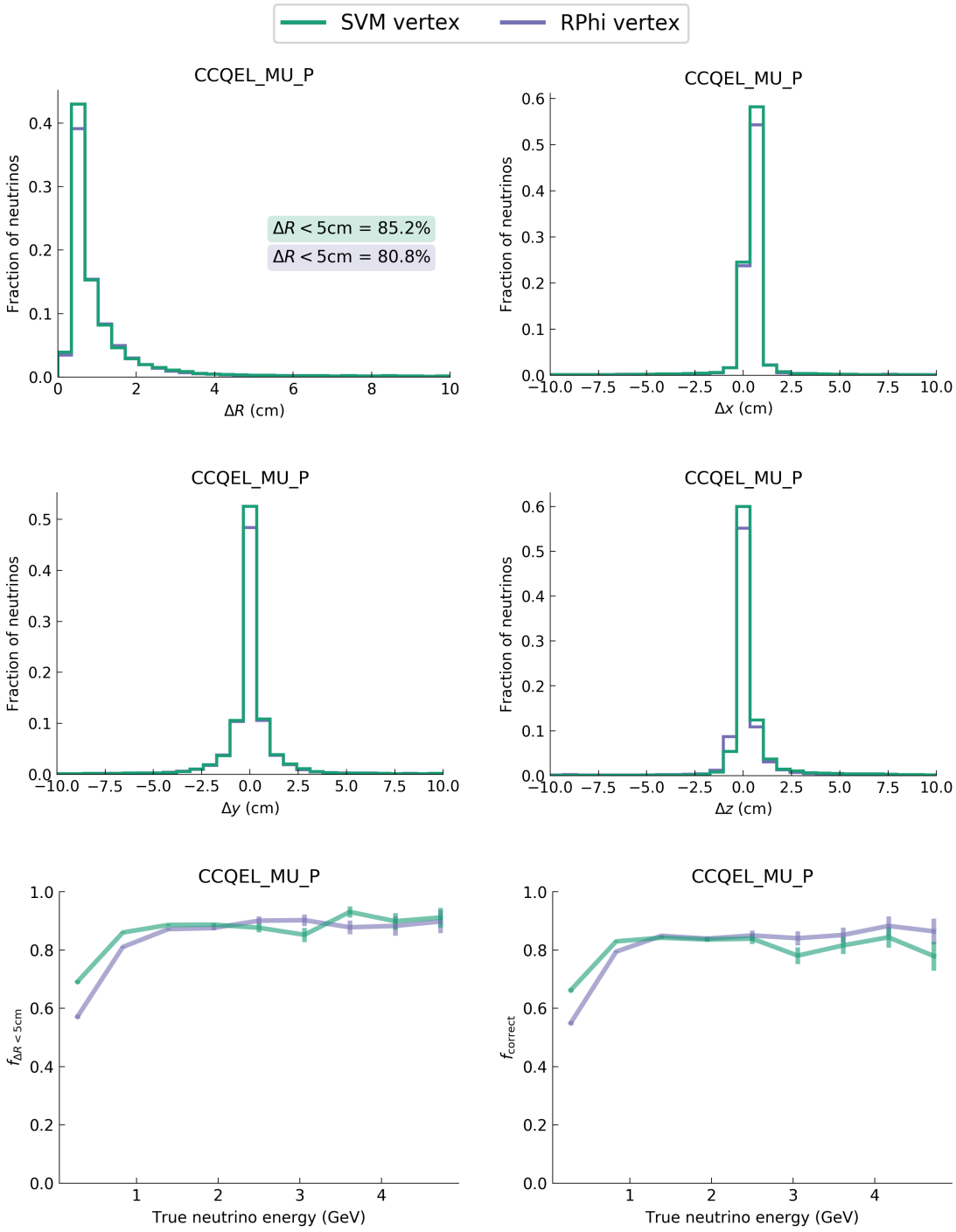


Figure A.19: Vertex reconstruction quality for the two algorithms for the BNB ν_μ CC quasi-elastic $\mu + 1p$ channel.

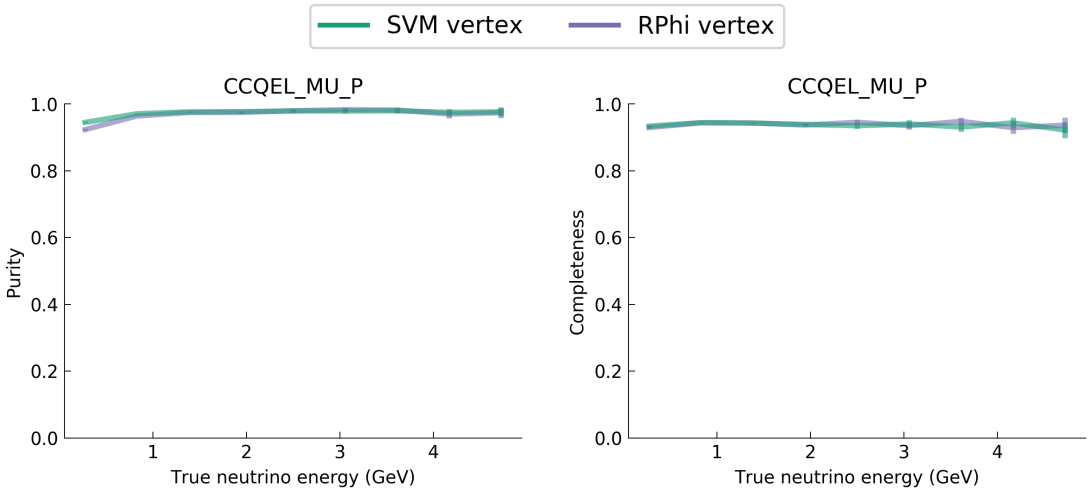


Figure A.20: MC particle matching quality for the two algorithms for the BNB ν_μ CC quasi-elastic $\mu + 1p$ channel.

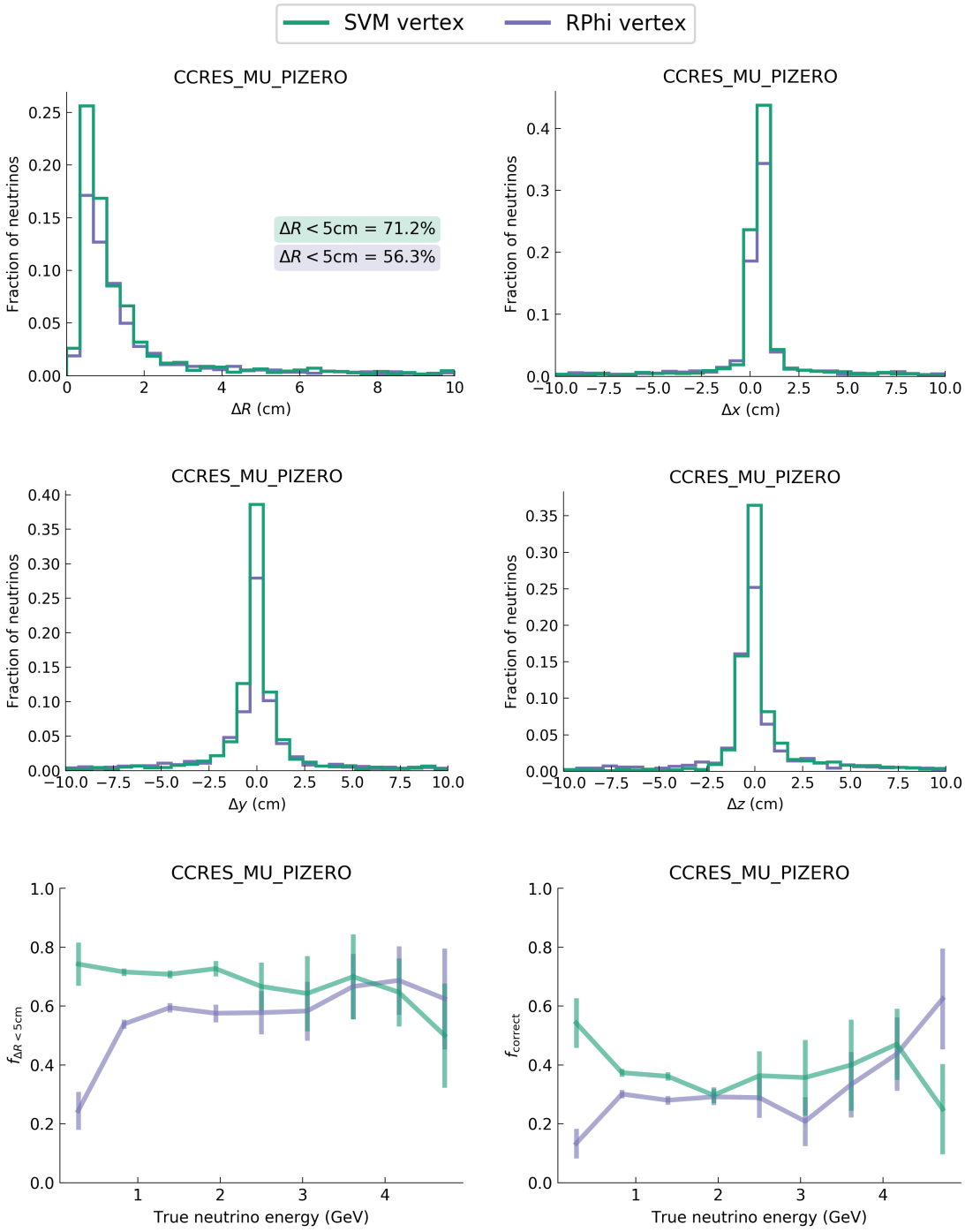


Figure A.21: Vertex reconstruction quality for the two algorithms for the BNB ν_μ CC resonant $\mu + \pi^0$ channel.

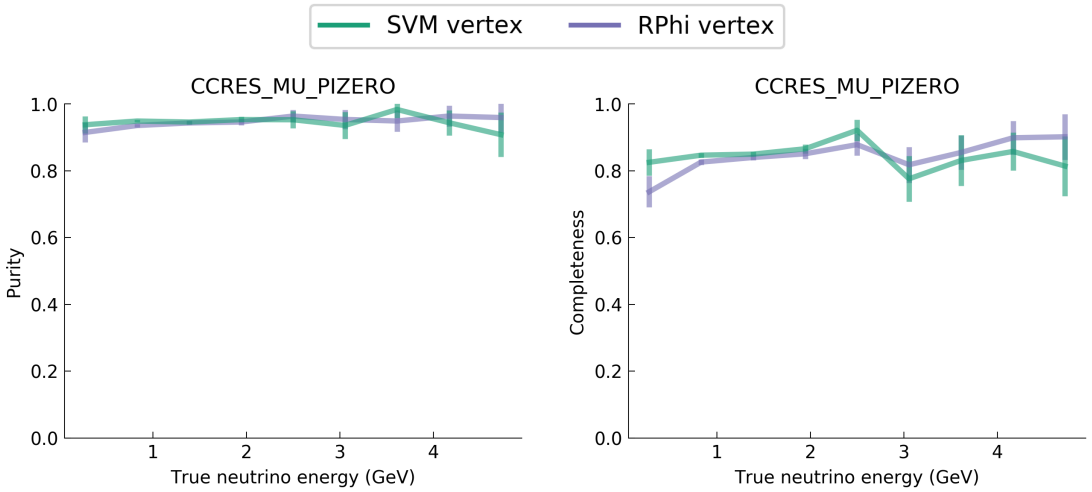


Figure A.22: MC particle matching quality for the two algorithms for the BNB ν_μ CC resonant $\mu + \pi^0$ channel.

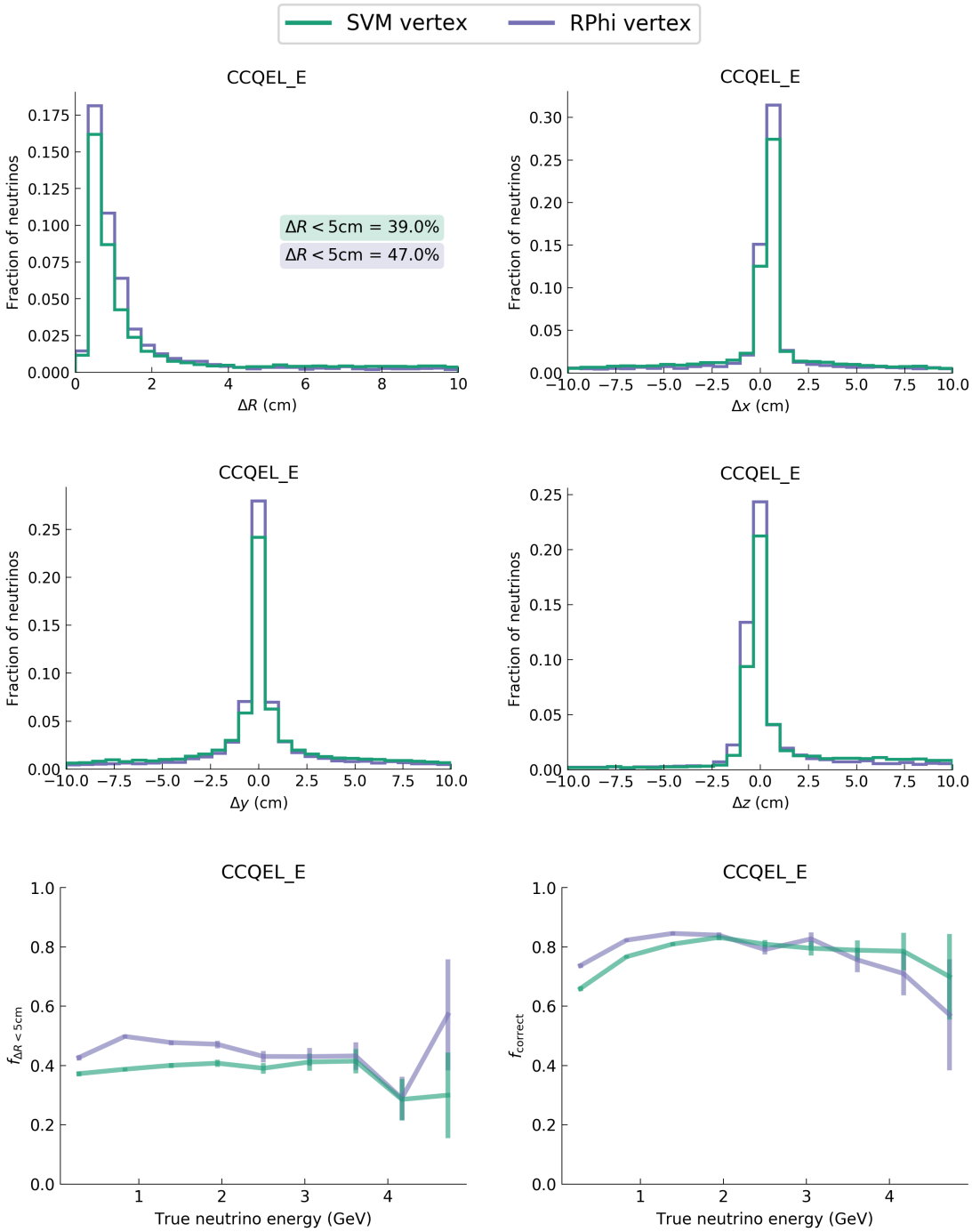


Figure A.23: Vertex reconstruction quality for the two algorithms for the BNB intrinsic ν_e CC quasi-elastic $e + 0p$ channel.

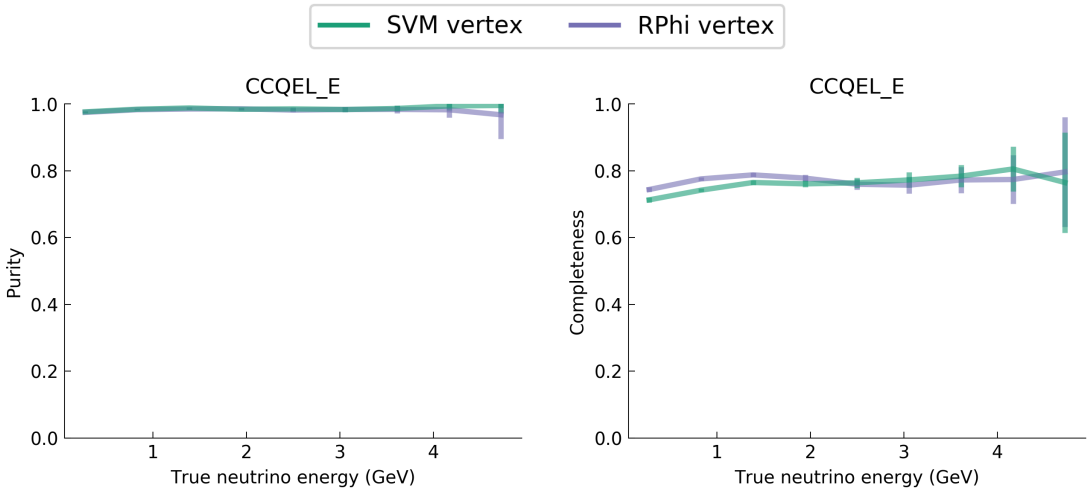


Figure A.24: MC particle matching quality for the two algorithms for the BNB intrinsic ν_e CC quasi-elastic $e + 0p$ channel.

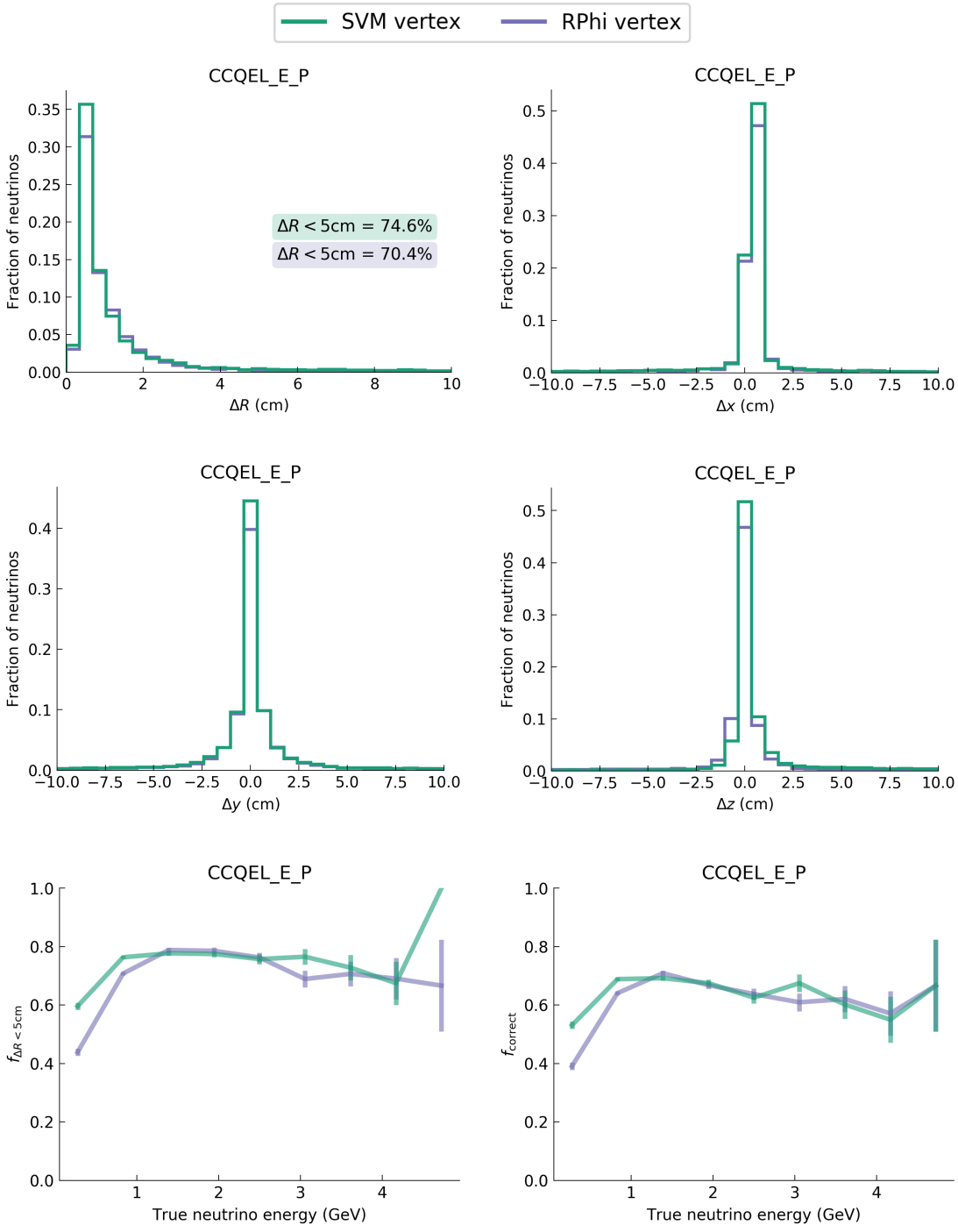


Figure A.25: Vertex reconstruction quality for the two algorithms for the BNB intrinsic ν_e CC quasi-elastic $e + 1p$ channel.

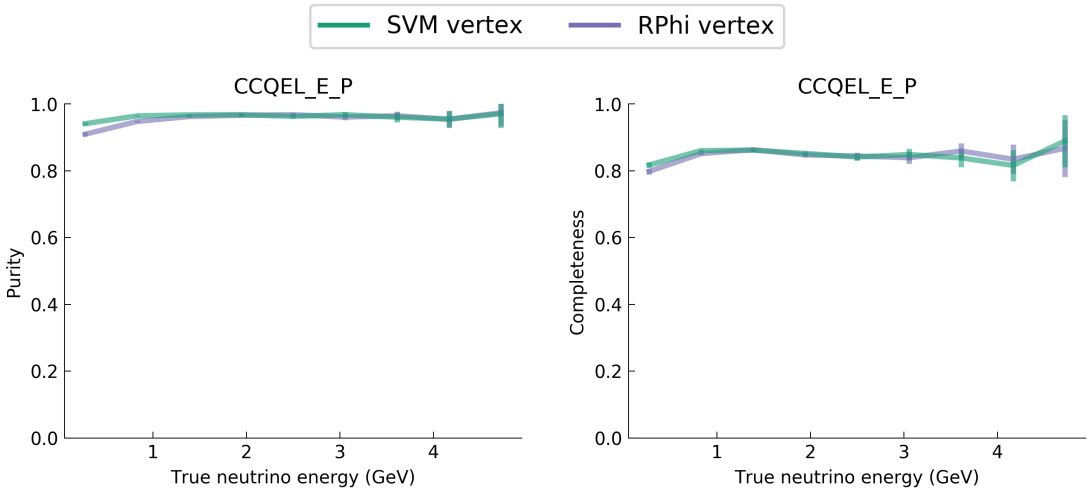


Figure A.26: MC particle matching quality for the two algorithms for the BNB intrinsic ν_e CC quasi-elastic $e + 1p$ channel.

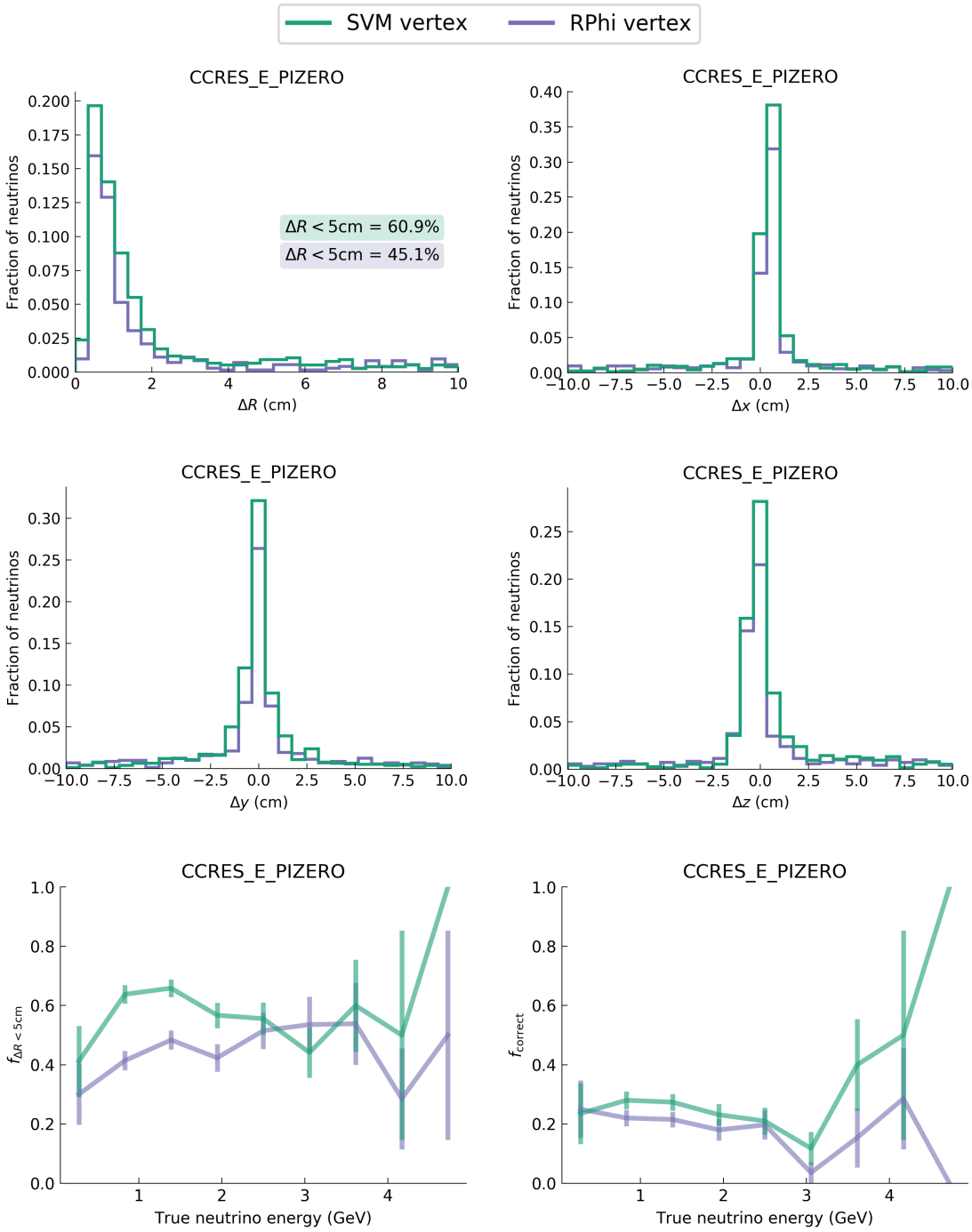


Figure A.27: Vertex reconstruction quality for the two algorithms for the BNB intrinsic ν_e CC resonant $e + \pi^0$ channel.

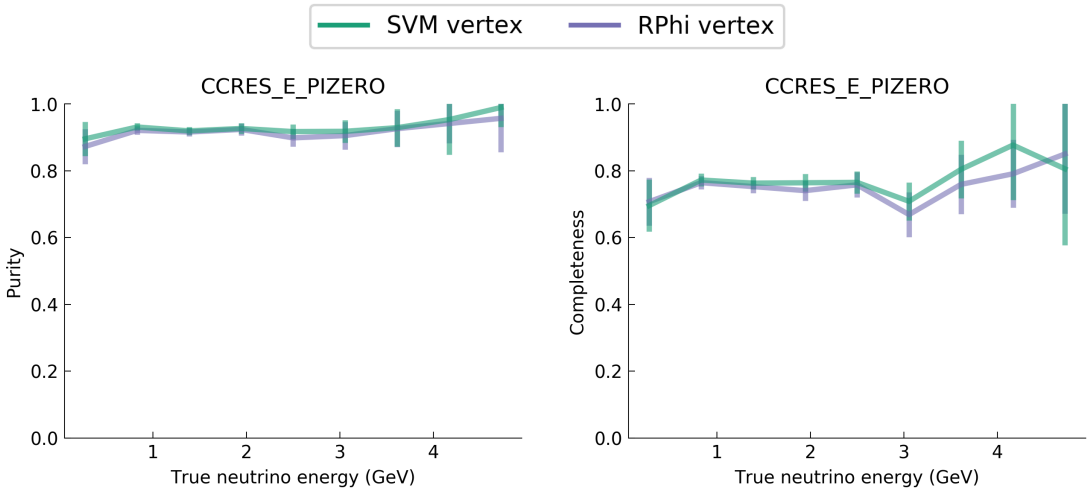


Figure A.28: MC particle matching quality for the two algorithms for the BNB intrinsic ν_e CC resonant $e + \pi^0$ channel.

Appendix B

FOMA-related mathematical derivations

B.1 Derivation of Equation 6.34

We begin by hijacking the derivation of Equation 6.28 at 6.24. Neglecting density effect corrections (for reasons discussed in the main text) and terms $\mathcal{O}(T'^2)$ and above yields

$$T'\frac{\mathrm{d}T'}{\mathrm{d}x} = -\frac{\xi'\chi}{2mc^2}. \tag{B.1}$$

This can be integrated directly, noting that the scaled modal kinetic energy $T'(x = R) = 0$ by definition of the particle range R, so

$$\int_0^{T'} t\mathrm{d}t = -\frac{\xi'\chi}{2mc^2}\int_R^x \mathrm{d}x' \tag{B.2}$$

$$\therefore T'^2 = \frac{\xi'\chi}{mc^2}(R-x) \tag{B.3}$$

$$T' = \pm\sqrt{\frac{\xi'\chi(R-x)}{mc^2}}. \tag{B.4}$$

Physically, we require $T' \in [0, \infty)$. This gives us the final expression for T':

$$T'(x; R, m) = \sqrt{\frac{\xi'\chi(R-x)}{mc^2}}. \tag{B.5}$$

Differentiating with respect to x yields

$$-\frac{\mathrm{d}T'}{\mathrm{d}x} = \frac{1}{2}\sqrt{\frac{\xi'\chi}{mc^2(R-x)}}. \tag{B.6}$$

Finally, using $-\mathrm{d}E/\mathrm{d}x = -\mathrm{d}T/\mathrm{d}x = -mc^2 \cdot \mathrm{d}T'/\mathrm{d}x$ (see Section 6.3.1),

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{1}{2}\sqrt{\frac{\xi'\chi mc^2}{R-x}}. \tag{B.7}$$

B.2 Second-order approximation to the mode

A justifiable extension to this work would be developing a method based on the equivalent second-order approximation to the mode. The second-order model is specified by Equations 6.28 and 6.45 and requires Newton's method to solve (see Section 6.3). To loosen the dependence of the method on ξ' and χ, we wish to proceed in analogy with the first-order method described in Section 7.5. However, due to the parametrization of the system by T' and its complex dependence on ξ' and χ, it is not possible to condense the parameters in the same way for the second-order approximation. Whatever method is chosen must require numerical inversion of either Equation 6.28 or Equation 6.3, and a straight line fit to estimate the mode as per Section 7.4. The method should then lead to the definition of a low-dimensional parameter space that can be used to separate the particles by mass.

One way to satisfy these requirements is to begin by inverting Equation 6.33 using Newton's method and the best estimate of the detector parameters:

$$0 = f\left(T'; \xi', \chi, \frac{\mathrm{d}E}{\mathrm{d}x}\right) := \frac{\mathrm{d}E}{\mathrm{d}x} + \xi'\left(\frac{(T'+1)^2}{T'(T'+2)}\left[2\ln\left(1+T'\right)+\chi\right]-1\right) \tag{B.8}$$

$$\therefore \frac{\partial f}{\partial T'} = -\frac{2\xi'(T'+1)\left[\chi - T'(T'+2) + 2\ln(1+T')\right]}{T'^2(T'+2)^2}. \tag{B.9}$$

This can be solved numerically for T' at each hit in the Bragg peak ($R - x < 10\,\mathrm{cm}$, as per Section 6.3.2) using Newton's method and the Modified Box model (Equation 6.43). Then, from Equation 6.28,

$$\frac{1}{3}\ln\left(1+\frac{2T'^2}{\chi}\right) + \sqrt{\frac{\chi}{2}}\arctan\left(\sqrt{\frac{2}{\chi}}T'\right) - T' = M'(m)(R-x), \tag{B.10}$$

where

$$M'(m) := \frac{2\xi'}{3mc^2}. \tag{B.11}$$

So a plot of the LHS (again using the best estimate for χ) against $(R-x)$ using repeated median regression should yield a straight line passing through the origin with gradient

$M'(m)$. However, in choosing inexact values for ξ' and χ, it is likely that the dependence on l is not completely accounted for – and also likely that discrepancies in the parameters and in the model will lead to a distortion and blurring of the predicted relationship. We therefore define a 2D parameter space of the gradient and residual intercept (as per our method in first order case) or a 3D parameter space also incorporating l. Similarly to the first-order method, training would be performed by populating the a graph with stopping particles of various types and identifying regions with each particle ID, either manually or using a clustering algorithm, on simulation or data. Once trained, the values of M and l for a given particle can then be mapped to a particle label.

Appendix C

PidNet architecture

The architectures of the two- and three-class classifiers used in Chapter 8, based on the autoencoder-learnt representation are described in Figure C.1 using the typical TensorFlow convention. The architecture of the autoencoder itself is given in Figure C.2.

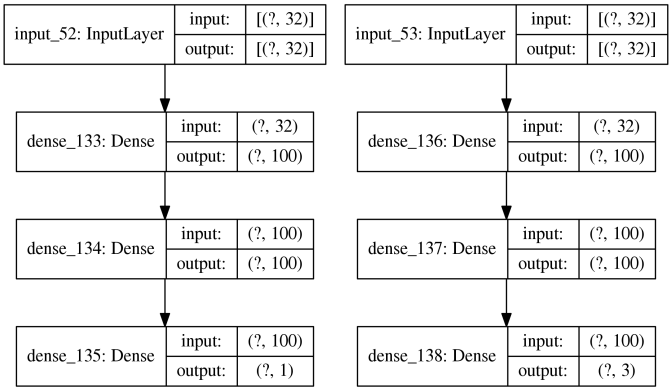


Figure C.1: The architectures of the binary (left) and three-class (right) classifiers used in this chapter.

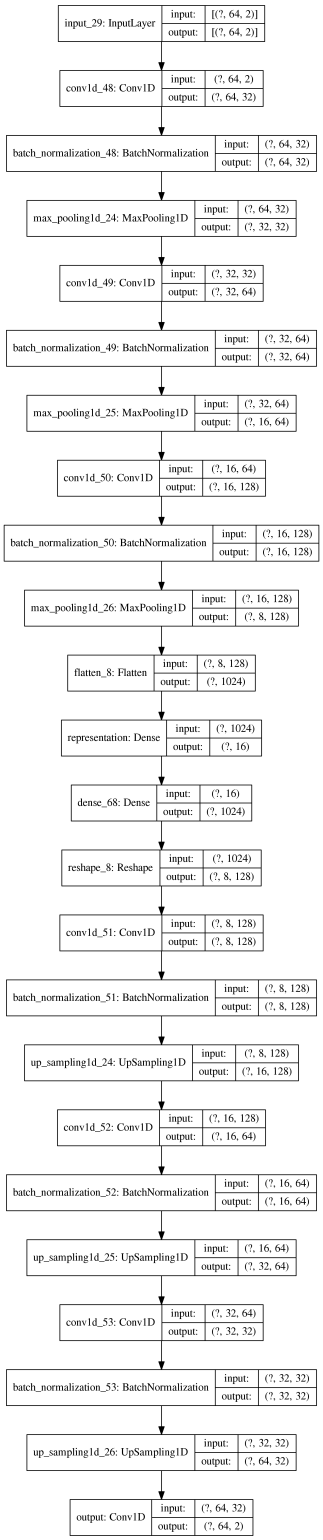


Figure C.2: The architecture of the convolutional autoencoder used in this chapter.

Bibliography

[1] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, 2012.

[2] A. Sushkov, W. Kim, D. Dalvit, and S. Lamoreaux. New experimental limits on non-Newtonian forces in the micrometer range. *Physical Review Letters*, 107(17):171101, 2011.

[3] G. Jungman, M. Kamionkowski, and K. Griest. Supersymmetric dark matter. *Physics Reports*, 267(5-6):195–373, 1996.

[4] F. Hasert, H. Faissner, W. D. Krenz, et al. Search for elastic muon-neutrino electron scattering. *Physics Letters. Section B*, 46(1):121–124, 1973.

[5] F. Hasert, S. Kabe, W. Krenz, et al. Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment. *Nuclear Physics B*, 73(1):1–22, 1974.

[6] C. Berger, W. Lackas, F. Raupach, et al. Jet analysis of the $v(9.46)$ decay into charged hadrons. *Physics Letters B*, 82(3-4):449–455, 1979.

[7] CDF Collaboration. Observation of top quark production in $p\ p$ collisions with the Collider Detector at Fermilab. *Physical Review Letters*, 74(14):2626, 1995.

[8] J.-E. Augustin, A. M. Boyarski, M. Breidenbach, et al. Discovery of a narrow resonance in $e^+\ e^-$ annihilation. *Physical Review Letters*, 33(23):1406, 1974.

[9] SNO Collaboration. Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory. *Physical Review Letters*, 89(1):011301, 2002.

[10] Super-Kamiokande Collaboration. Evidence for oscillation of atmospheric neutrinos. *Physical Review Letters*, 81(8):1562, 1998.

[11] M. Gell-Mann, P. Ramond, R. Slansky, P. Van Nieuwenhuizen, and D. Freedman. Supergravity. 1979.

[12] R. N. Mohapatra and G. Senjanović. Neutrino mass and spontaneous parity nonconservation. *Physical Review Letters*, 44(14):912, 1980.

[13] Particle Data Group. Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2020(8):083C01, 2020.

[14] MiniBooNE Collaboration. Significant excess of electronlike events in the MiniBooNE short-baseline neutrino experiment. *Physical Review Letters*, 121:221801, 2018. doi:10.1103/PhysRevLett.121.221801. URL `https://link.aps.org/doi/10.1103/PhysRevLett.121.221801`.

[15] C. Athanassopoulos, L. Auerbach, D. Bauer, et al. Candidate events in a search for $\nu_\mu \to \nu_e$ oscillations. *Physical Review Letters*, 75(14):2650, 1995.

[16] F. Varanini. ICARUS detector: present and future. In *EPJ Web of Conferences*, volume 164, page 07017. EDP Sciences, 2017.

[17] C. Adams, C. Andreopoulos, J. Asaadi, et al. LAr1-ND: testing neutrino anomalies with multiple LArTPC detectors at Fermilab. *arXiv preprint arXiv:1309.7987*, 2013.

[18] Hyper-Kamiokande Collaboration. Letter of intent: The Hyper-Kamiokande Experiment - detector design and physics potential. *arXiv preprint arXiv:1109.3262*, 2011.

[19] R. Gandhi, B. Kayser, M. Masud, and S. Prakash. The impact of sterile neutrinos on CP measurements at long baselines. *Journal of High Energy Physics*, 2015(11):39, 2015.

[20] R. Acciarri, C. Adams, R. An, et al. A proposal for a three detector short-baseline neutrino oscillation program in the Fermilab booster neutrino beam. *arXiv preprint arXiv:1503.01520*, 2015.

[21] MicroBooNE Collaboration. Design and construction of the MicroBooNE detector. *Journal of Instrumentation*, 12(02):P02017, 2017.

[22] H. Becquerel. The radio-activity of matter. *Nature*, 63(1634):396, 1901.

[23] W. Röntgen. Weitere Beobachtungen über die Eigenschaften der X-Strahlen. *Annalen der Physik*, 300(1):18–37, 1898.

[24] E. Rutherford. VIII. Uranium radiation and the electrical conduction produced by it. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 47(284):109–163, 1899.

[25] J. Chadwick. Intensitätsverteilung im magnetischen Spectrum der β-Strahlen von radium B+ C. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 16:383, 1914.

[26] W. Pauli. Dear radioactive ladies and gentlemen. *Physics Today*, 31:27, 1930.

[27] J. Chadwick. Possible existence of a neutron. *Nature*, 129(3252):312–312, 1932.

[28] I. Curie. Un nouveau type de radioactivité. *Comptes rendus des seances de l'Academie des sciences (CRAS)*, 198:254–256, 1934.

[29] E. Fermi. Tentativo di una teoria dei raggi β. *Il Nuovo Cimento (1924-1942)*, 11(1):1, 1934.

[30] H. Bethe and R. Peierls. The "neutrino". *Nature*, 133(3362):532–532, 1934.

[31] C. Cowan, F. Reines, F. Harrison, and H. Kruse. Detection of the free neutrino: a confirmation. *Science*, 124(3212):103–104, 1956.

[32] C.-S. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Physical Review*, 105(4):1413, 1957.

[33] G. Danby, J. M. Gaillard, K. Goulianos, et al. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Physical Review Letters*, 9(1):36, 1962.

[34] M. L. Perl, G. Abrams, A. Boyarski, et al. Evidence for anomalous lepton production in e^+ - e^- annihilation. *Physical Review Letters*, 35(22):1489, 1975.

[35] DONUT Collaboration. Observation of tau neutrino interactions. *Physics Letters B*, 504(3):218–224, 2001.

[36] S. L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961.

[37] A. Salam and J. C. Ward. Electromagnetic and weak interactions. In *Selected Papers Of Abdus Salam (with commentary)*, pages 210–213. World Scientific, 1994.

[38] S. Weinberg. A model of leptons. *Physical Review Letters*, 19(21):1264, 1967.

[39] O. Chibani. New algorithms for the Vavilov distribution calculation and the corresponding energy loss sampling. *IEEE Transactions on Nuclear Science*, 45(5):2288–2292, 1998.

[40] UA2 Collaboration. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $\overline{p}p$ collider. *Physics Letters B*, 122(CERN-EP-83-25):476–485, 1983.

[41] K. Lande, R. Davis, B. Cleveland, et al. Results from the Homestake solar neutrino observatory. In *High Energy Physics*, volume 900802, pages 667–675. 1991.

[42] SNO Collaboration. Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory. *Physical Review Letters*, 87(7):071301, 2001.

[43] S. Weinberg. *The quantum theory of fields*, volume 2. Cambridge University Press, 1995.

[44] C. Giunti and C. W. Kim. *Fundamentals of neutrino physics and astrophysics*. Oxford University Press, 2007.

[45] T.-P. Cheng, L.-F. Li, and T.-P. Cheng. Gauge theory of elementary particle physics. 1984.

[46] J. F. Donoghue, E. Golowich, and B. R. Holstein. *Dynamics of the Standard Model*, volume 35. Cambridge University Press, 2014.

[47] F. Boehm, P. Vogel, and P. Vogel. *Physics of massive neutrinos*. Cambridge University Press, 1992.

[48] B. Kayser. On the quantum mechanics of neutrino oscillation. *Physical Review D*, 24(1):110, 1981.

[49] Z. Maki, M. Nakagawa, and S. Sakata. Remarks on the unified model of elementary particles. *Progress of Theoretical Physics*, 28(5):870–880, 1962.

[50] A. Pilaftsis. CP violation and baryogenesis due to heavy Majorana neutrinos. *Physical Review D*, 56(9):5431, 1997.

[51] M. Flanz, E. A. Paschos, and U. Sarkar. Baryogenesis from a lepton asymmetric universe. *Physics Letters B*, 345(3):248–252, 1995.

[52] G. Fogli, E. Lisi, A. Marrone, and D. Montanino. Status of atmospheric neutrino $\nu_\mu \to \nu_\tau$ oscillations and decoherence after the first K2K spectral data. *Physical Review D*, 67(9):093006, 2003.

[53] E. Lisi, A. Marrone, and D. Montanino. Probing possible decoherence effects in atmospheric neutrino oscillations. *Physical Review Letters*, 85(6):1166, 2000.

[54] V. Barger, J. Learned, P. Lipari, et al. Neutrino decay and atmospheric neutrinos. *Physics Letters B*, 462(1-2):109–114, 1999.

[55] V. Barger, J. Learned, S. Pakvasa, and T. J. Weiler. Neutrino decay as an explanation of atmospheric neutrino observations. *Physical Review Letters*, 82(13):2640, 1999.

[56] C. Spiering. Towards high-energy neutrino astronomy. In *From Ultra Rays to Astroparticles*, pages 231–263. Springer, 2012.

[57] N. G. Cooper. Los Alamos Science, number 25 (1997): Celebrating the neutrino. Technical report, Los Alamos National Lab., NM (United States), 1997.

[58] IceCube Collaboration. Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert. *Science*, 361(6398):147–151, 2018.

[59] G. Mangano, G. Miele, S. Pastor, et al. Relic neutrino decoupling including flavour oscillations. *Nuclear Physics B*, 729(1-2):221–234, 2005.

[60] KamLAND Collaboration. Experimental investigation of geologically produced antineutrinos with kamland. *Nature*, 436(7050):499–503, 2005.

[61] W. C. Haxton, R. Hamish Robertson, and A. M. Serenelli. Solar neutrinos: status and prospects. *Annual Review of Astronomy and Astrophysics*, 51:21–61, 2013.

[62] L. Anchordoqui, T. Paul, S. Reucroft, and J. Swain. Ultrahigh energy cosmic rays: The state of the art before the auger observatory. *International Journal of Modern Physics A*, 18(13):2229–2366, 2003.

[63] A. Bernstein, Y.-f. Wang, G. Gratta, and T. West. Nuclear reactor safeguards and monitoring with antineutrino detectors. *Journal of Applied Physics*, 91(7):4672–4676, 2002.

[64] P. Adamson, K. Anderson, M. Andrews, et al. The NuMI neutrino beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 806:279–306, 2016.

[65] F. Suekane, T. Iwamoto, H. Ogawa, O. Tajima, and H. Watanabe. An overview of the KamLAND 1-kiloton liquid scintillator. *arXiv preprint physics/0404071*, 2004.

[66] SAGE Collaboration. Measurement of the solar neutrino capture rate in SAGE. *Nuclear Physics B-Proceedings Supplements*, 118:39–46, 2003.

[67] GALLEX Collaboration. GALLEX solar neutrino observations: Results for GALLEX IV. *Physics Letters B*, 447(1-2):127–133, 1999.

[68] D. A. Glaser. Some effects of ionizing radiation on the formation of bubbles in liquids. *Physical Review*, 87(4):665, 1952.

[69] P. A. Cherenkov. Visible emission of clean liquids by action of γ radiation. In *Doklady Akademii Nauk SSSR*, volume 2, pages 451–454. 1934.

[70] SNO Collaboration. The Sudbury neutrino observatory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 449(1-2):172–207, 2000.

[71] MiniBooNE Collaboration. The MiniBooNE detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 599(1):28–46, 2009.

[72] Super-Kamiokande Collaboration. Calibration of the Super-Kamiokande detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 737:253–272, 2014.

[73] IceCube Collaboration. First year performance of the IceCube neutrino telescope. *Astroparticle Physics*, 26(3):155–173, 2006.

[74] MINOS Collaboration. First observations of separated atmospheric ν_μ and anti-ν_μ events in the MINOS detector. *Phys. Rev. D*, 73:072002, 2006. doi:10.1103/PhysRevD.73.072002. `hep-ex/0512036`.

[75] ICARUS Collaboration. Underground operation of the ICARUS T600 LAr-TPC: first results. *Journal of Instrumentation*, 6(07):P07011, 2011.

[76] MicroBooNE Collaboration. Design and construction of the MicroBooNE detector. *Journal of Instrumentation*, 12(02):P02017, 2017.

[77] ProtoDUNE Collaboration. The single-phase ProtoDUNE technical design report. *arXiv preprint arXiv:1706.07081*, 2017.

[78] J. Detwiler. Recent results from KamLAND. *arXiv preprint hep-ex/0311007*, 2003.

[79] KamLAND Collaboration. Constraints on θ_{13} from a three-flavor oscillation analysis of reactor antineutrinos at KamLAND. *Physical Review D*, 83(5):052002, 2011.

[80] KamLAND Collaboration. Precision measurement of neutrino oscillation parameters with KamLAND. *Physical Review Letters*, 100(22):221803, 2008.

[81] Super-Kamiokande public site: Event Display. URL `http://www-sk.icrr.u-tokyo.ac.jp/sk/detector/eventdisplay-e.html`.

[82] Super-Kamiokande Collaboration. Measurements of the atmospheric neutrino flux by Super-Kamiokande: energy spectra, geomagnetic effects, and solar modulation. *Physical Review D*, 94(5):052001, 2016.

[83] T2K public site: About T2K, 2013. URL `http://t2k-experiment.org/t2k/`.

[84] T2K Collaboration. Combined analysis of neutrino and antineutrino oscillations at T2K. *Physical Review Letters*, 118(15):151801, 2017.

[85] T2K Collaboration. Constraint on the matter-antimatter symmetry-violating phase in neutrino oscillations. *arXiv preprint arXiv:1910.03887*, 2019.

[86] MINOS Collaboration. The magnetized steel and scintillator calorimeters of the MINOS experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 596(2):190–228, 2008.

[87] MINOS+ Collaboration. New results from MINOS and MINOS+. *IOP Conference Series: Journal of Physics*, 888(1):012017, 2017.

[88] MINOS Collaboration. Precision measurement of the speed of propagation of neutrinos using the MINOS detectors. *Physical Review D*, 92(5):052005, 2015.

[89] MINOS+ Collaboration. Search for sterile neutrinos in MINOS and MINOS+ using a two-detector fit. *Physical Review Letters*, 122(9):091803, 2019.

[90] Daya Bay and MINOS+ Collaborations. Improved constraints on sterile neutrino mixing from disappearance searches in the MINOS, MINOS+, Daya Bay, and Bugey-3 experiments. *Physical Review Letters*, 125(7):071801, 2020.

[91] SNO public site: Some interesting SNO events, 2003. URL `https://www.sno.phy.queensu.ca/sno/events/`.

[92] SNO Collaboration. The Sudbury Neutrino Observatory. *Nuclear Physics B*, 908:30–51, 2016.

[93] Daya Bay Collaboration. Latest results from Daya Bay. In *Journal of Physics: Conference Series*, volume 873, page 012031. IOP Publishing, 2017.

[94] Daya Bay Collaboration. Measurement of the electron antineutrino oscillation with 1958 days of operation at Daya Bay. *Physical Review Letters*, 121(24):241805, 2018.

[95] Daya Bay Collaboration. Improved search for a light sterile neutrino with the full configuration of the Daya Bay experiment. *Physical Review Letters*, 117(15):151802, 2016.

[96] KATRIN Collaboration. The KATRIN neutrino mass experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 623(1):442–444, 2010.

[97] SNO+ Collaboration. Neutrinoless double beta decay with SNO+. In *Journal of Physics: Conference Series*, volume 375, page 042015. IOP Publishing, 2012.

[98] R. Arnold, C. Augier, J. Baker, et al. Probing new physics models of neutrinoless double beta decay with SuperNEMO. *The European Physical Journal C*, 70(4):927–943, 2010.

[99] J. J. Gómez Cadenas, V. Álvarez, F. Borges, et al. Present status and future perspectives of the NEXT experiment. *Advances in High Energy Physics*, 2014, 2014.

[100] LSND Collaboration. Evidence for neutrino oscillations from the observation of anti-neutrino (electron) appearance in an anti-neutrino (muon) beam. *Physical Review*, 64(hep-ex/0104049):112007, 2001.

[101] MiniBooNE Collaboration. The neutrino flux prediction at MiniBooNE. *arXiv preprint arXiv:0806.1449*, 2008.

[102] M. Sorel, J. M. Conrad, and M. H. Shaevitz. Combined analysis of short-baseline neutrino experiments in the (3+1) and (3+2) sterile neutrino oscillation hypotheses. *Physical Review D*, 70(7):073004, 2004.

[103] H. Paes, S. Pakvasa, and T. J. Weiler. Sterile-active neutrino oscillations and shortcuts in the extra dimension. *Physical Review D*, 72(9):095017, 2005.

[104] T. Goldman, G. Stephenson Jr, and B. McKellar. Multichannel oscillations and relations between LSND, KARMEN, and MiniBooNE, with and without CP violation. *Physical Review D*, 75(9):091301, 2007.

[105] S. Gninenko. MiniBooNE anomaly and heavy neutrino decay. *Physical Review Letters*, 103(24):241802, 2009.

[106] T. Katori, V. A. Kostelecký, and R. Tayloe. Global three-parameter model for neutrino oscillations using Lorentz violation. *Physical Review D*, 74(10):105009, 2006.

[107] J. A. Harvey, C. T. Hill, and R. J. Hill. Standard model gauging of the Wess-Zumino-Witten term: Anomalies, global currents, and pseudo-Chern-Simons interactions. *Physical Review D*, 77(8):085017, 2008.

[108] V. Shiltsev. Fermilab proton accelerator complex status and improvement plans. *Modern Physics Letters A*, 32(16):1730012, 2017.

[109] J. de Vries. *Identifying Muon Neutrino Charged-Current Interactions in the MicroBooNE Detector*. Ph.D. thesis, University of Cambridge, 2019.

[110] MicroBooNE Collaboration. Booster neutrino flux prediction at MicroBooNE. 13:10, 2018.

[111] P. Adamson, K. Anderson, M. Andrews, et al. The NuMI neutrino beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 806:279–306, 2016.

[112] S. R. Soleti. *Search for a low-energy excess of electron neutrinos in MicroBooNE.* Ph.D. thesis, University of Oxford, 2019.

[113] V. Meddage. Electron attenuation measurement using cosmic ray muons at the MicroBooNE LArTPC. *arXiv preprint arXiv:1710.00396*, 2017.

[114] F. Cavanna, A. Ereditato, and B. Fleming. Advances in liquid argon detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 907:1–8, 2018.

[115] H. Chen, K. Chen, G. De Geronimo, et al. Readout electronics for the MicroBooNE LAr TPC, with CMOS front end at 89K. *Journal of Instrumentation*, 7(12):C12004, 2012.

[116] M. Auger, M. Del Tutto, A. Ereditato, et al. A novel cosmic ray tagger system for liquid argon TPC neutrino detectors. *Instruments*, 1(1):2, 2017.

[117] MicroBooNE Collaboration. Design and construction of the MicroBooNE cosmic ray tagger system. *Journal of Instrumentation*, 14(04):P04004, 2019.

[118] MicroBooNE Collaboration. Noise characterization and filtering in the MicroBooNE liquid argon TPC. *Journal of Instrumentation*, 12(08):P08003, 2017.

[119] MicroBooNE Collaboration. Ionization electron signal processing in single phase LArTPCs. part II: Data/simulation comparison and performance in MicroBooNE. *Journal of Instrumentation*, 13(07):P07007, 2018.

[120] C. Andreopoulos, A. Bell, D. Bhattacharya, et al. The GENIE neutrino Monte Carlo generator. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 614(1):87–104, 2010.

[121] GEANT4 Collaboration. GEANT4 — a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[122] D. Heck, J. Capdevielle, G. Schatz, T. Thouw, et al. CORSIKA: A Monte Carlo code to simulate extensive air showers," report fzka 6019, forschungszentrum karlsruhe. 1998.

[123] R. Brun, L. Urban, F. Carminati, et al. GEANT: detector description and simulation tool. Technical report, CERN, 1993.

[124] E. D. Church. LArSoft: a software package for liquid argon time projection drift chambers. *arXiv preprint arXiv:1311.6774*, 2013.

[125] J. Marshall and M. Thomson. The Pandora software development kit for pattern recognition. *The European Physical Journal C*, 75(9):439, 2015.

[126] MicroBooNE Collaboration. The Pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the MicroBooNE detector. *The European Physical Journal C*, 78(1):82, 2018.

[127] MicroBooNE Collaboration. Comparison of ν_μ-Ar multiplicity distributions observed by MicroBooNE to GENIE model predictions. *The European Physical Journal C*, 79(3):248, 2019.

[128] L. Landau. On the energy loss of fast particles by ionization. 1944.

[129] P. Vavilov. Ionization losses of high-energy heavy particles. *Soviet Physics – JETP*, 5, 1957.

[130] H. A. Bethe. Passage of radiations through matter. *Experimental Nuclear Physics*, 1953.

[131] R. M. Sternheimer. The density effect for the ionization loss in various materials. *Physical Review*, 88(4):851, 1952.

[132] J. Weston. `bethe-faster`: A framework for particle identification and energy estimation using a sequential Monte Carlo method. URL `http://github.com/jackanth/bethe-faster`.

[133] H. Bichsel. Straggling in thin silicon detectors. *Reviews of Modern Physics*, 60(3):663, 1988.

[134] M. Lauge. dE/dx *Measurements in the ATLAS Detector*. Niels Bohr Institute, Copenhagen University, 2017.

[135] R. Sternheimer, M. Berger, and S. M. Seltzer. Density effect for the ionization loss of charged particles in various substances. *Atomic Data and Nuclear Data Tables*, 30(2):261–271, 1984.

[136] MicroBooNE Collaboration. Detector calibration using through going and stopping muons in the MicroBooNE LArTPC. Technical report, MicroBooNE Technical Note, 2018.

[137] MicroBooNE Collaboration. Calibration of the charge and energy response of the MicroBooNE liquid argon time projection chamber using muons and protons. *arXiv preprint arXiv:1907.11736*, 2019.

[138] M. Mooney. The MicroBooNE experiment and the impact of space charge effects. *arXiv preprint arXiv:1511.01563*, 2015.

[139] MicroBooNE Collaboration. A measurement of the attenuation of drifting electrons in the MicroBooNE LArTPC, 2017.

[140] ArgoNeuT Collaboration. A study of electron recombination using highly ionizing particles in the ArgoNeuT Liquid Argon TPC. *Journal of Instrumentation*, 8(08):P08005, 2013.

[141] MicroBooNE Collaboration. Calibration of the charge and energy loss per unit length of the MicroBooNE liquid argon time projection chamber using muons and protons. *Journal of Instrumentation*, 15(03):P03022, 2020.

[142] S. Burdin, M. Horbatsch, and W. Taylor. A correction to Birks' Law in liquid argon ionization chamber simulations for highly ionizing particles. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 664(1):111–116, 2012.

[143] MicroBooNE Collaboration. First muon-neutrino charged-current inclusive differential cross section measurement for MicroBooNE run 1 data. 2018. doi: 10.2172/1573222.

[144] MicroBooNE Collaboration. Ionization electron signal processing in single phase LArTPCs. part i: Algorithm description and quantitative evaluation with MicroBooNE simulation. *Journal of Instrumentation*, 13(07):P07006, 2018.

[145] MicroBooNE Collaboration. MCC9 overlay GENIE simulated BNB and cosmic data. Technical report, MicroBooNE Technical Note, 2019.

[146] H.-J. Yang, B. P. Roe, and J. Zhu. Studies of boosted decision trees for MiniBooNE particle identification. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 555(1-2):370–385, 2005.

[147] MicroBooNE Collaboration. Deep neural network for pixel-level electromagnetic particle identification in the microboone liquid argon time projection chamber. *Physical Review D*, 99(9):092001, 2019.

[148] S. Alonso-Monsalve and L. H. Whitehead. Image-based model parameter optimization using model-assisted generative adversarial networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2020.

[149] W. Yin, K. Kann, M. Yu, and H. Schütze. Comparative study of CNN and RNN for natural language processing. *arXiv preprint arXiv:1702.01923*, 2017.

[150] H.-Y. Lee, H.-Y. Tseng, J.-B. Huang, M. Singh, and M.-H. Yang. Diverse image-to-image translation via disentangled representations. In *Proceedings of the European conference on computer vision (ECCV)*, pages 35–51. 2018.

[151] F. Locatello, S. Bauer, M. Lucic, et al. Challenging common assumptions in the unsupervised learning of disentangled representations. In *international conference on machine learning*, pages 4114–4124. 2019.

[152] E. Alpaydin. *Introduction to machine learning*. MIT press, 2020.

[153] X. Guo, X. Liu, E. Zhu, and J. Yin. Deep clustering with convolutional autoencoders. In *International conference on neural information processing*, pages 373–382. Springer, 2017.

[154] G. E. Hinton and S. T. Roweis. Stochastic neighbor embedding. In *Advances in neural information processing systems*, pages 857–864. 2003.

[155] A. Vaswani, N. Shazeer, N. Parmar, et al. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008. 2017.

[156] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[157] A. Radford, J. Wu, R. Child, et al. Language models are unsupervised multitask learners. *OpenAI Blog*, 1(8):9, 2019.

[158] A. v. d. Oord, Y. Li, and O. Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

[159] D. P. Kingma and M. Welling. An introduction to variational autoencoders. *arXiv preprint arXiv:1906.02691*, 2019.

[160] ArgoNeuT Collaboration. First observation of low energy electron neutrinos in a liquid argon time projection chamber. *Physical Review D*, 95(7):072005, 2017.

[161] MicroBooNE Collaboration. Michel electron reconstruction using cosmic-ray data from the MicroBooNE LArTPC. *Journal of Instrumentation*, 12(09):P09014, 2017.

[162] E. Shibamura, A. Hitachi, T. Doke, et al. Drift velocities of electrons, saturation characteristics of ionization and W-values for conversion electrons in liquid argon, liquid argon-gas mixtures and liquid xenon. *Nuclear Instruments and Methods*, 131(2):249–258, 1975.

[163] S. Das. A simple alternative to the Crystal Ball function. *arXiv preprint arXiv:1603.08591*, 2016.

[164] C.-Y. Yi and H.-S. Han. A Monte Carlo algorithm for the Vavilov distribution. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 149(3):263–271, 1999.

[165] P. J. Rousseeuw and C. Croux. Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424):1273–1283, 1993.

[166] DUNE Collaboration. Long-baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) conceptual design report. Volume 4: The DUNE detectors at LBNF. *arXiv preprint arXiv:1601.02984*, 2016.

[167] DUNE public site: Home, 2020. URL `https://www.dunescience.org/`.