FASER tracker performance and electron energy resolution in the FASERν pilot detector

John W. Spencer

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Shih-Chieh Hsu, Chair

David B. Kaplan

Henry J. Lubatti

Program Authorized to Offer Degree:
Department of Physics

University of Washington

Abstract

FASER tracker performance and electron energy resolution in the FASERν pilot detector

John W. Spencer

Chair of the Supervisory Committee:
Professor Shih-Chieh Hsu
Physics

In collider experiments, very light particles are produced in the far-forward direction with small angle relative to the beam axis and can travel a long distance before decaying. The ForwArd Search ExpeRiment (FASER) is aptly located 480 m downstream from the ATLAS interaction point where background is minimal, so optimally positioned to detect very light new particles. Such particles are likely to decay in FASER to two particles, so two-track vertexing can be used to detect these particles.

There is an unexplored energy region 350 GeV - 6 TeV in which the neutrino-nucleon charged current cross-section is not yet known. The FASERν emulsion detector, positioned just upstream of FASER, will detect collider-produced neutrinos for the very first time. In addition, the interface detector enables track matching between the FASER spectrometer and the FASERν emulsion detector, which enables separate cross section measurements for mu neutrinos and antineutrinos. For electron neutrinos, the outgoing electron will initiate an electromagnetic shower, whose profile characteristics will be used to reconstruct the electron's energy.

The FASERν pilot detector was installed in the TI18 maintenance tunnel in 2018 and collected $12.2\,\mathrm{fb}^{-1}$ of data, mainly from muons originating at the ATLAS IP. These muons initiate knock-on electron EM showers. By clustering these showers and reconstructing their energies, the muon energy spectrum can be validated, as well as the procedure for

reconstructing ν_e energy in FASERν.

This thesis studies performance at simulation level as well as data analysis. The FASER tracker performance is studied. For the FASERν emulsion detector, the electron seeding and clustering algorithms are motivated and developed, and the electron energy resolution is presented. For the FASERν pilot run, the validated electron clustering algorithm is run on the data samples in the last 29 layers of the FASERν pilot detector and the results are compared with the equivalent simulation samples.

TABLE OF CONTENTS

Page

List of Figures iv

List of Tables x

Chapter 1: Introduction 1

Chapter 2: Theory background 4
2.1 The Standard Model 4
2.2 Neutrino Interactions 7
2.3 Electron interactions 10
2.4 Muon interactions 11

Chapter 3: Experimental Apparatus 13
3.1 The Large Hadron Collider 13
3.2 The FASER Spectrometer 14
3.3 The Interface Tracker 21
3.4 The FASERν Emulsion Detector 21
3.5 The Pilot Detector 27

Chapter 4: Tracker and Emulsion Datasets 29
4.1 Tracker Performance Dataset 29
4.2 Pilot Data Sample 29
4.3 Pilot Detector MC Simulation Sample 31

Chapter 5: Tracking and Reconstruction with FASER 34
5.1 Tracking in FASER 34
5.2 Estimation of tracker performance 40

Chapter 6: Electron Clustering and Energy Reconstruction 42
6.1 Muon Subtraction 42
6.2 Electron clustering algorithm 46
6.3 Electron energy regression 50

Chapter 7: Pilot Run Analysis 53
7.1 Comparison of raw basetrack parameters between MC and data 55
7.2 Muon identification 55
7.3 Electron clustering sample preparation 58
7.4 Cluster purification in the pilot data 62
7.5 Electron energy regression with the pilot data 63
7.6 Comparison of clustered electron showers between MC and data 67
7.7 Electron clustering 72

Chapter 8: Results and Conclusion 74

Chapter 9: Outlook 76

Bibliography 78

Appendix A: Helix Tracker Reconstruction using FaserMC 82

Appendix B: Neutrino Signal / Background Signatures 87
B.1 Electron neutrino CC events 87
B.2 Muon neutrino CC events 90
B.3 Tau neutrino CC events 90
B.4 Background Processes 91

Appendix C: Truth-Leve Simulation Study of Electromagnetic Showers 98
C.1 Characteristics of Electromagnetic Showers 98
C.2 Electron Energy Reconstruction with Total Track Count 99
C.3 Electron Energy Reconstruction with Maximum Track Depth and Multiplicity 100
C.4 Electron Energy Reconstruction with the Full Shower Profile 102

Appendix D: Tracker and Tracker+IFT Performance 106
D.1 Tracker performance 109

Appendix E: IFT Track Matching between FASER Tracker and FASERν Emulsion Detector 112

LIST OF FIGURES

Figure Number — Page

1.1 Existing measurements of the neutrino-nucleon charged-current cross-sections for the three flavours of neutrino (ν_e, ν_μ, ν_τ), along with the expected incident neutrino spectra on FASERν. 2

2.1 Feynman diagrams for neutrino-nucleon neutral current (left) and charged current (right) interactions. 8

2.2 One possible Feynman diagram for neutrino-induced pion production. 8

2.3 Feynman diagrams for dominant production mechanisms for ν_e, ν_μ, and ν_τ. . 10

2.4 Feynman diagrams for electron bremsstrahlung. 10

2.5 Schematic illustration of a muon-induced knock-on electron production event. 11

2.6 Feynman diagram for a ν_τ CC event producing a τ which promptly decays to a pion. 12

3.1 Schematic diagram of the Large Hadron Collider, as well as the four major detectors ATLAS, CMS, ALICE, and LHCb positioned along the collider ring and the infrastructure to produce the high-energy proton beams. 13

3.2 Schematic diagram of the region downstream of the ATLAS IP. The vertical axis is scaled disproportionately for clarity. Charged particles are deflected by the LHC magnets while neutral particles continue on a straight line until they are absorbed by the TAN neutral particle absorber. Neutrinos, dark photons, and other new physics particles may continue uninterrupted, travelling approximately 480 m until they reach FASER. 15

3.3 Schematic diagram illustrating the placement and orientation of FASER relative to the collider ring of the LHC. 16

3.4 Computer-generated model of the FASER spectrometer (red), FASERν emulsion detector (green), IFT (white), and calorimeter (purple). The tracking stations are indicated with red arrows. The masses of each detector element are displayed in parentheses. 16

3.5 FASER spectrometer installed in the TI12 maintenance tunnel of the LHC for commissioning as of 5 February 2021. 17

3.6 (left) SCT module; (right) Eight SCT modules assembled into one FASER tracking plane. 18
3.7 One of the three SCT tracking stations assembled from three tracking planes. 19
3.8 (far-left) Rectangular section of an emulsion film; (center-left) schematic diagram of emulsion film showing two emulsion layers on either side of a 200 μm plastic base; (center-right) electron microscope photograph of AgBr crystals in emulsion film; (far-right) image of the emulsion track left by a secondary particle emerging from a 10 GeV π. 22
3.9 Schematic diagram of support structure that will contain the FASERν emulsion detector. 23
3.10 Rendering of the placement of the FASER spectrometer and the FASERν emulsion detector in the TI12 tunnel. The ATLAS IP is off-screen to the right. 24
3.11 The two tungsten (W) and lead (Pb) modules of the FASERν pilot detector placed in TI12 in September-October 2018. 27

4.1 Energy spectra of μ^+ (red) and μ^- (blue) arriving at the FASERν pilot detector. 32
4.2 Spectra of $\tan\theta_x$ and $\tan\theta_y$ for μ^+ (red) and μ^- (blue) arriving at the FASERν pilot detector. 32

5.1 Bitstream format for readout from the SCT modules in the FASER spectrometer. 35
5.2 Schematic diagram of hits being combined into hit clusters on one side of a given SCT module. The 3 green hit strips are combined into a single hit cluster, as are the two red hit strips and the single blue hit strip. 36
5.3 Schematic diagram of hit clusters on opposite sides of a given SCT module to produce a space-point. 37
5.4 Schematic diagram of space points being used for track-finding and track-fitting using a Kalman filter. 37
5.5 Schematic of helix track parameters. 38

6.1 Illustration of four basetracks (nodes A-D) for step 5. The links between basetracks are deleted in the sequence shown at right. 43
6.2 Visualization of the position (above) and angular (below) triangles whose areas are used to evaluate the credibility of a link. 44
6.3 Example illustration of the track-linking procedure. After all suitable basetrack links are made (left), the procedure to remove multiple links is applied and the most credible path is selected (middle). The straight tracks are then reconnected (right). 44

6.4 Diagram showing the electron clustering procedure, starting with the seed basetrack (blue), constructing the cone (green), selecting basetracks within the cone, extrapolating and applying cuts on Δr and $\Delta\theta$, and adding basetracks to the electron cluster (black). Basetracks which fail the matching criteria (red) are not clustered. 47

6.5 Example reconstruction of electron cluster with background contamination from other clusters (left), and clustered electron after purification (right). . . 48

6.6 Plot of projections of clustered basetracks to emulsion film of seed basetrack. 49

6.7 Event display of multiple reconstructed electron clusters purified with IP cut. 50

6.8 Plots of total number of clustered basetracks N_{total} (left) and maximum clustered basetrack multiplicity in a single layer N_{max} (right) as functions of energy. 51

6.9 Plot of the depth (emulsion plate index) of the shower maximum iz_{max} as a function of energy. 52

7.1 (top-left) Histogram of x for emulsion hits in truth MC simulation superposed with x for basetracks in one 2×2 subvolume block of the pilot run data samples. (rop-right) Histogram of y for emulsion hits in truth MC simulation superposed with y for basetracks in one 2×2 subvolume block of the pilot run data samples. (bottom-left) Histogram of $\tan\theta_x$ for emulsion hits in truth MC simulation superposed with $\tan\theta_x$ for basetracks in pilot run data samples. (bottom-right) Histogram of $\tan\theta_y$ for emulsion hits in truth MC simulation superposed with $\tan\theta_y$ for basetracks in pilot run data samples. 56

7.2 Two-dimensional event display showing xy- (left), xz- (middle), and yz-projections (right) of all basetracks within $(200\,\mu\text{m})^2$ region centered on $(x, y) = (50, 50)$ mm. The lengths of the basetracks has been extrapolated to 1 mm to show the relationship between basetracks in successive layers. 59

7.3 Two-dimensional event display showing xy- (left), xz- (middle), and yz-projections (right) of all basetracks that are components of reconstructed straight tracks within $(200\,\mu\text{m})^2$ region centered on $(x, y) = (50, 50)$ mm. 59

7.4 Two-dimensional event display showing xy- (left), xz- (middle), and yz-projections (right) of all basetracks that are components of reconstructed straight tracks within $(200\,\mu\text{m})^2$ region centered on $(x, y) = (50, 50)$ mm and that have passed the further cuts to be classified as penetrating tracks. 59

7.5 Two-dimensional event display showing xy- (left), xz- (middle), and yz-projections (right) of all basetracks within $(200\,\mu\text{m})^2$ region centered on $(x, y) = (50, 50)$ mm that are not part of penetrating straight tracks. 60

7.6 Event display of all clustered electrons within $(20\,\text{mm})^2$ area centered on $(x, y) = (50, 50)$ mm. 61

7.7 Distribution of impact parameters for clustered basetracks belonging to the same truth cluster as the seed basetrack (orange) and clustered basetracks belonging to a different truth cluster from the seed basetrack (blue). 62
7.8 Scatter plot of number of clustered e^+/e^- basetracks as a function of the truth electron energy (blue points) along with a fit (orange line). Note that the number of basetracks grows approximately linearly with energy. 63
7.9 Electron energy resolution for one-dimensional energy regression using the total number of clustered, purified basetracks N_{total}. Note that below $E_{\text{truth}} \approx 300\,\text{GeV}$ the resolution is very poor and the results unreliable. 64
7.10 Comparison of number of truth basetracks per layer (orange) and number of clustered basetracks per layer (blue). 64
7.11 Distribution of muon energy losses (left) and distribution of energies of produced knock-on electrons (right) in the pilot run simulation. Note that the electron energy distribution is shifted to the left as compared with the muon energy loss distribution. 65
7.12 Histogram of opening angle between knock-on electron and parent muon for electrons with energies $E > 10\,\text{GeV}$ (blue) and $E > 100\,\text{GeV}$ (orange). 66
7.13 Overlapping histograms showing the comparison of reconstructed electron energies in the GEANT4 MC simulation (blue) and in the FASERnu pilot run data (orange). 66
7.14 Scatter plot of actual angular spectra (θ_x, θ_y) of reconstructed basetracks in the pilot run. 69
7.15 Comparison of histograms of $\tan\theta_x$ and $\tan\theta_y$ for simulation (solid color) and pilot data (error bars). 70
7.16 Comparison of histograms of x and y for simulation (solid color) and pilot data (error bars). The separation is due to spreading of simulation hits from inside the subvolume to outside, not balanced by spreading from outside the subvolume to inside (as with the pilot detector). 70
7.17 Comparison of histograms of iz for simulation (solid color) and pilot data (error bars). The inconsistency does not affect the data/MC comparison, see main text. 71
7.18 OpenGL GEANT4 event display of a muon (blue) which initiates a delta-ray (green) at a vertex (red). 71
7.19 Sample of truth knock-on electron clusters (left) and reconstructed electron clusters (right) from MC sample. 72
7.20 Number of clustered electron showers as a function of plate index iz (left) and of reconstructed electron energy (right). 73

A.1 A schematic diagram showing the hits registered by a μ^+ in the FASER spectrometer as well as the hybrid track that the particle takes while travelling through it. The tracking stations are shown in grey, the hits are shown in blue, and the direction of the nonzero magnetic field in the dipole volumes between the tracking stations is shown in magenta. The black track segments are linear while the coloured segments are helical. 83

A.2 Illustration of a helix and its projection to the plane transverse to its axis, along with the track radius R, the initial transverse position (x_0, y_0), the initial azimuthal angle ϕ_0, and the helix pitch λ. 84

A.3 (left) p_T resolution of FASER spectrometer using the helix tracker and the legacy MC simulation; (right) p_T resolution of FASER spectrometer using the improved hybrid tracker. 85

A.4 Resolutions for the helix fitting parameters (top-left) x_0, (top-right) y_0, (bottom-left) ϕ_0, and (bottom-right) λ using the hybrid helix/linear tracker and the legacy MC simulation. Data points indicate resolution for tracks reconstructed without the IFT (red) and with the IFT (black). 86

B.1 Lepton signatures for each of the three neutrino-nucleon CC interactions. . . 88

B.2 Feynman diagram for a simplified electromagnetic shower. The incident electron emits a photon, which splits into an electron-positron pair, each of which emit further photons. 89

B.3 Feynman diagram of electron neutrino CC event, which initiates a hadronic shower by stimulating emission of a virtual gluon which decays into a neutral pion. 89

B.4 Expected distribution of background positive (red) and negative (blue) muons incident on FASERν. 92

B.5 Energy spectra for emerging photons (left) and e^+/e^- (right) emerging from incident μ^-. 93

B.6 Energy spectra for emerging neutrons (top-left), antineutrons (top-right), K_S (bottom-left), and K_L (bottom-right) emerging from incident μ^-. 94

B.7 Energy spectra for emerging π^+ (top-left), and π^- (top-right), K^+ (bottom-left), and K^- (bottom-right) emerging from incident μ^-. 95

B.8 Table of expected background induced by muons and signal particles passing through FASERν. 96

B.9 Depth profile for 10^4 electrons (left) and negative muons (right) at 1 TeV energy and at normal incidence on FASERν. 96

B.10 Signal-background separation for 10^4 electrons (left) and negative muons (right) at 1 TeV energy with cuts $E > 100$ MeV, $\tan\theta < 0.05$, and $x^2 + y^2 < (200\,\mu\text{m})^2$. 97

C.1 Plots of the total number of tracks as a function of electron energy along with the fitted linear regression. 99

C.2 Plots of the maximum track multiplicity in a single layer (left), and the depth of the shower maximum (right), as a function of electron energy. A linear regression is fitted to the track multiplicity, while a logarithmic regression is fitted to the depth of the shower maximum. 101

D.1 Plot of the charge misidentification rate as a function of incident muon energy. 107

D.2 Tracker x position resolution (top-left), y position resolution (top-right) and fractional momentum resolution (bottom-left) versus truth energy, with fitted curves. 110

E.1 CAD model of the FASERν emulsion detector and the IFT of the FASER spectrometer, with distance relevant for track matching shown in green. . . . 112

E.2 $\epsilon_{\text{vertexID}}$ as a function of incident neutrino energy. 114

E.3 Muon neutrino-antineutrino CC cross-section sensitivity as a function of energy. Error bars show statistical (dark blue) and statistical+systematic (light blue) uncertainties. 115

LIST OF TABLES

Table Number Page

2.1 Particle content of the SM. 5

2.2 List of common Standard Model decay modes which may contribute to neutrino production. 9

3.1 Estimated number of neutrino/antineutrino CC interactions in FASERν for each of the three flavours of neutrino. [27] 25

ACKNOWLEDGMENTS

The author wishes to express sincere appreciation to the University of Washington and his doctoral research advisor Shih-Chieh Hsu and the entire FASER/FASERν collaboration. He would like to thank his collaborators for their contributions to the work presented in this thesis. In particular, he thanks Aki Ariga and Tomoko Ariga for their support with FEDRA and the pilot data samples, Tomohiro Inada for his support on emulsion reconstruction, Jeffrey Gao, Van Tha Bik Lian, Yen-Ting Chou, and Ali Salman for their debugging help, Tobias Boeckh and Ke Li for their input on the tracking offline software. His work was supported in part by Heising-Simons Foundation Grant Nos. 2020-1840 and Simons Foundation Grant No. 623683, as well as by teaching assistantships and associateships provided by the University of Washington Physics Department.

DEDICATION

to my parents, Loan and William Spencer, without whom this thesis would not have been possible, and to Ann Nelson who instilled in me a love of particle physics which continues to this day

Chapter 1

INTRODUCTION

The neutrino is widely regarded in physics as the most elusive particle in the universe. Its mean free path in lead is one light-year ($\sim 10^{13}$ km), and billions pass through every person every second without their noticing. There are three varieties, or *flavours,* of neutrino, and while it is known that their interactions with "ordinary" matter are extremely weak, there are large gaps in the data where the theoretically predictions of interaction likelihood have not yet been confirmed by experiment.

The strength of an interaction is best measured by a quantity in particle physics known as the *cross-section.* The muon neutrino cross sections at low energy have been determined by beam dump experiments, and at high energy by cosmic neutrinos, but in the 400 GeV$-$1 TeV energy range, the muon neutrino cross sections have not been measured.

Because of the low theoretical prediction for their cross sections, neutrinos are particularly difficult to detect. In particular, colliders are the dominant production source for neutrinos in the unmeasured energy range specified above. Furthermore, other particles which may exist beyond the Standard Model of particle physics, but which are also very light, may be detectable by an experimental apparatus not yet realized at the Large Hadron Collider.

The ForwArd Search ExpeRiment, or FASER [1], is an experiment which aims to detect these weakly-interacting particles. It is optimally positioned to receive these particles, upstream from the ATLAS interaction point and tangent to the collider ring. An additional detector component, FASERν [2], will use an entirely different detector specialized to detecting and distinguishing neutrino interactions, although it does not have the capability to distinguish neutrinos from antineutrinos on its own. In conjunction with one another, and an additional interface tracker linking FASER with FASERν, this combined detector will be

able to measure the cross sections of mu neutrinos and antineutrinos separately.

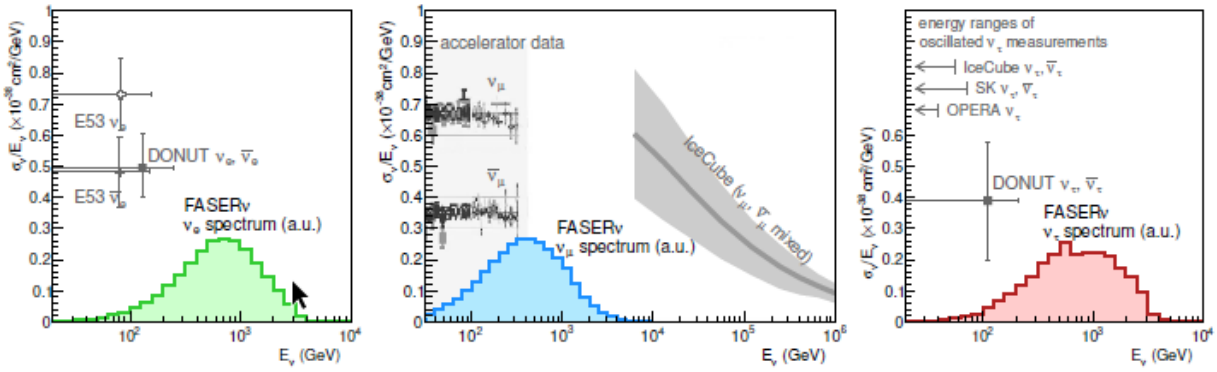

Figure 1.1: Existing measurements of the neutrino-nucleon charged-current cross-sections for the three flavours of neutrino (ν_e, ν_μ, ν_τ), along with the expected incident neutrino spectra on FASERν.

Figure 1.1 shows the existing neutrino-nucleon charged-current cross-section measurements that have been made using existing detectors.[2] In particular, the cross-sections for the electron neutrino have been measured by the E53 dump experiment [3]; muon neutrinos and antineutrinos have been separately measured at low energy by beam-dump experiments, and at high energy by the IceCube detector at the South Pole [4]; and tau neutrinos have been directly observed by the Direct Observation of Nu Tau (DONuT) experiment [5]. Oscillated tau neutrinos have also been observed in experiments such as OPERA [6] and Super Kamiokande [7]; the energy ranges of these observations are indicated in the right-most panel of Figure 1.1.

The author of this thesis has personally contributed to a preliminary study of the muon-type neutrino-antineutrino charged-current cross-section sensitivity, as well as an evaluation of the FASER tracker performance and an algorithm to reconstruct electrons and their energies in the FASERν detector, which was developed by studying real data from the FASERν pilot detector and its simulation. The tracker performance and electron energy

regression will be discussed in this thesis because of their relevance to studying neutrino cross-sections with FASERν: The tracker performance is essential for reconstructing the tracks of muons in the FASER spectrometer, which will be used to study muon neutrinos and antineutrinos, whereas the electron energy reconstruction approach is a necessary step for reconstructing the energies of electron neutrinos and antineutrinos.

This thesis is outlined as follows. In Chapter 2 we present the theoretical background knowledge for elementary particle physics, including the Standard Model and neutrino interactions. In Chapter 3 the Large Hadron Collider is described, along with the existing and prospective collider experiments, including FASER and FASERν. In Chapter 4 the simulated and real data sets used for analysis are discussed. In Chapter 4 the tracker reconstruction algorithm is discussed and the performance results presented. In Chapter 6 the electron clustering algorithm for the MC simulation is described and performance results presented. In Chapter 7 the pilot run analysis is carried out and the data and MC are compared to demonstrate their agreement. In Chapter 8 the outlook of future work with FASER and FASERν is discussed.

Several studies done by the author are relegated to the Appendices. Appendix A discusses the tracker performance using a legacy simulation of the FASER spectrometer. Appendix B discusses characteristics of the signal neutrino events in the emulsion detector as well as background processes for each of these signals. Appendix C discusses a truth-level study of electron energy and its dependence on several parameters associated with truth electron showers. Appendix D discusses the charge misidentification in the FASER spectrometer. Appendix E introduces the concept of track matching between FASER and FASERν.

Chapter 2
THEORY BACKGROUND

In this chapter, the Standard Model of particle physics is introduced. The particle content is presented, and the strong and weak interactions are then discussed. The neutrino is introduced, and the characteristic neutrino-nucleon neutral-current and charged-current processes are outlined. Finally, physics which cannot be explained by the Standard Model is briefly described, and additional postulated particles are mentioned.

2.1 The Standard Model

The Standard Model of particle physics, was originally proposed by Sheldon Glashow to describe the unification of the electromagnetic (EM) and weak interactions, and expanded upon by Abdus Salam and Steven Weinberg in 1967 to describe the Higgs mechanism [8]. Presently, neutrino masses and oscillations [9] remain the only verified phenomena observed to contradict the Standard Model.

The Standard Model consists of twelve particles with half-integer spin (known as fermions) and five particles of integer spin (known as bosons), as shown in Table 2.1. Fermions comprise all matter content in the observable universe, whereas bosons mediate three of the four fundamental forces of nature – electromagnetism (EM), the strong nuclear force, and the weak nuclear force, to the exclusion of gravity (for which there has yet to be formulated a quantum theory and the state-of-the-art macroscopic description is Einstein's general theory of relativity). These three interactions are mediated by the massless *photon* (γ), the massless *gluon* (g), and the massive *weak bosons* ($Z^0, W^{\pm}$), respectively. These particles have spin 1, and correspondingly are referred to as *vector bosons*. In contrast, the final particle of the Standard Model has spin 0 and is the famous *Higgs boson* (H), which does not mediate a

Table 2.1: Particle content of the SM.

SM	Fermions			Bosons	
Quarks	u	c	t	$W^{\pm}$	H
	d	s	b	Z^0	
Leptons	e	μ	τ	γ	
	ν_e	ν_μ	ν_τ	g	

force but rather is responsible for the masses of all particles in the Standard Model via the Higgs mechanism which will be discussed shortly.

According to modern quantum field theory, the Standard Model describes quantum fields that are invariant under local transformations by members of a continuous group – these are known as *gauge fields,* and their excitations or quanta give rise to the gauge bosons described above.

The fermions can be divided into quarks (the upper two rows) and leptons (the lower two rows). They are distinguished by their electric charge and the forces by which they interact. In particular, quarks may interact strongly, electromagnetically, or weakly, whereas the leptons interact only electromagnetically or weakly.

The first row of fermions contains the up-type quarks, which have a fractional electric charge of $+2/3$. The second row describes the down-type quarks, which have a fractional electric charge of $-1/3$. The respective up- and down-type quarks in the first generation are the *up* (u) and *down* (d) quarks, those in the second generation are the *charm* (c) and *strange* (s) quarks, and those in the third generation are the *top* (t) and *bottom* (b) quarks.

Quarks do not exist in isolation, but rather must exchange gluons to form strongly-interacting composite particles, also known as *hadrons*; this is because the potential energy due to the strong force between quarks increases with distance (in contrast to the electrostatic potential energy, which falls of as $\sim 1/r$ with separation distance r). As such, it is more energetically favourable to have a small quark separation distance. This phenomenon is

known as *confinement.* In particular, the quarks of the first generation exist to form protons and neutrons, which make up most of the visible matter in the universe.

The third row of table 2.1 describes the charged leptons, of which the electron is the most familiar member, which interact via the weak and EM forces due to their charge. The final row describes the neutral leptons, or neutrinos, which are electrically neutral, interact only weakly. Like the quarks, the leptons are also categorized into three generations. The respective charged leptons and neutrinos in the first generation are the *electron* (e) and the *electron neutrino* (ν_e), those in the second generation are the *muon* (μ) and the *mu neutrino* (ν_μ), and those in the third generation are the *tau particle* (τ) and the *tau neutrino* (ν_τ). However, unlike quarks, the leptons do not interact via the strong force and therefore do not form composite particles.

In addition, for each of the aforementioned particles, there exists a corresponding antiparticle of the same mass but opposite electric charge. Note that some particles are their own antiparticles, namely g, H, γ, and Z^0. Thus the inclusion of antiparticles does not precisely double the particle content of the Standard Model.

The weak interaction is parity-violating in that it affects only left-handed particles. The left- and right-handed projection operators are given by

$$L = \frac{1-\gamma_5}{2} \qquad \text{and} \qquad R = \frac{1+\gamma_5}{2}, \tag{2.1}$$

and are idempotent ($L^2 = L, R^2 = R$) and mutually orthogonal ($LR = RL = 0$).

In contrast to the strong interactions, the weak interactions can violate the discrete symmetries C, P, and CP. These discrete symmetry violations are critically important for *baryogenesis,* which seeks to explain the observation that the universe is overwhelmingly made of matter.

A Z boson can couple to up-type or down-type quarks. However, it is an experimental observation that the quarks must be of the same generation; in other words, there are no *flavour-changing neutral currents.*

In contrast, by exchanging W bosons, up-type quarks may couple to down-type quarks.

This may happen across multiple generations – for example, the u quark may couple to d, s, or b – but the process is less likely the more generations lie between the quarks. This is quantified by the Cabibbo-Kobayashi-Maskawa matrix [9] which describes the quark couplings to the W. Weak processes involving quarks separated by 0, 1, and 2 generations are known as *Cabibbo-favoured, Cabibbo-suppressed,* and *doubly Cabbibo-suppressed,* respectively.

The W boson may also couple neutrinos to charged leptons. However, unlike the quark sector, the charged leptons must couple to neutrinos of the same generation. This observation is important because the identification of neutrino flavour can and will be done by identification of the charged lepton produced in such an interaction.

2.2 Neutrino Interactions

The neutrino was first postulated by Wolfgang Pauli in 1930 [10] as a third particle in the beta decay of a radioactive nucleus, necessary for the energy spectrum of the daughter particles to be continuous. The name *neutrino* was suggested by Enrico Fermi, meaning "little neutral one." The particle was first observed experimentally in 1956 by the Cowan-Reines experiment [11].

The electron neutrino ν_e, the muon neutrino ν_μ, and the tau neutrino ν_τ are the most weakly-interacting in the Standard Model. They interact with the other particles of the SM in only two ways:

1. The neutrino may exchange a neutral Z^0 boson. In this case the incoming and outgoing particles are both neutrinos and have the same (zero) electric charge. For this reason the process is known as a *neutral current* or NC event.

2. The neutrino may exchange a charged W^+ or W^- boson. In this case the outgoing particle is a charged lepton e, μ, τ, whose lepton flavor corresponds to the incident neutrino and whose charge is anticorrelated with that of the W boson exchanged. Thus the electric charges of the incoming and outgoing particles are different. For this reason the process is known as a *charged current* or CC event.

The capability for FASERν to study neutral-current neutrino interactions has been discussed in [12]. In this thesis the emphasis will be on the charged-current interactions.

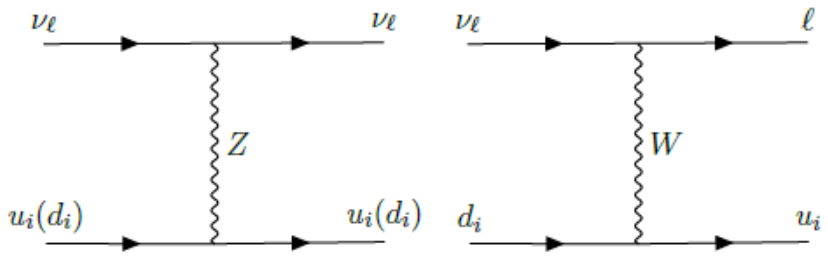

Figure 2.1: Feynman diagrams for neutrino-nucleon neutral current (left) and charged current (right) interactions.

Figure 2.1 shows the two types of neutrino-nucleon interactions. Here the nucleon is represented by a single quark; the other two quarks are spectators and do not participate in the interaction. Note that a sufficiently energetic incident neutrino can lead to neutrino pion production as shown in Figure 2.2. Neutral pions will decay electromagnetically, producing photons or electron-positron pairs.

Table 2.2 provides an abbreviated list of the most common SM hadron decay modes which may produce neutrinos, adapted from [2]. In particular, electron neutrinos are dominantly produced via semileptonic kaon decays and similar charged current processes with flavoured

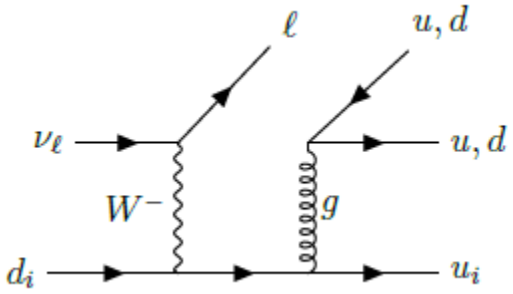

Figure 2.2: One possible Feynman diagram for neutrino-induced pion production.

Table 2.2: List of common Standard Model decay modes which may contribute to neutrino production.

Type	**Particles**	**Main Decays**
Pions	π^+	$\pi^+ \to \mu\nu_\mu$
Kaons	K^+, K_S, K_L	$K^+ \to \mu\nu_\mu, K \to \pi\ell\nu$
Baryons	$\Lambda, \Sigma^+, \Sigma^0, \Sigma^-, \Xi^0, \Xi^-, \Omega^-$	$\Lambda \to p\ell\nu$
Charmed Hadrons	$D^+, D^0, D_s^+, \Lambda_c^0, \Xi_c^+, \Xi_c^0$	$D \to K\ell\nu, D_s \to \tau\nu_\tau, \Lambda_c \to \Lambda\ell\nu$
Bottom Hadrons	$B^+, B^0, B_s^0, \Lambda_b^-, \Xi_b^0, \Xi_b^-$	$B \to D\ell\nu, \Lambda_b \to \Lambda_c\ell\nu$

baryons, while muon neutrinos are dominantly produced via leptonic charged kaon and pion decays, and tau neutrinos are dominantly produced via the leptonic decay of the D_s meson.

In addition to the neutral- and charged-current events described above, neutrinos can also undergo flavour oscillations – that is, a ν_e may oscillate into a ν_μ, and either of these may become a ν_τ, or vice versa. These are known as *neutrino oscillations,* and they arise due to the fact that the flavour eigenstates $|\nu_e\rangle, |\nu_\mu\rangle$, and $|\nu_\tau\rangle$ are not exactly the same as the mass eigenstates $|\nu_1\rangle, |\nu_2\rangle$, and $|\nu_3\rangle$. The mixing between the neutrino mass and flavour eigenstates is governed by the Pontecorvo-Maki-Nakagawa-Sakata matrix, analogous to the CKM matrix for quarks.

Strictly speaking, neutrino mixing and oscillation is not accounted for by the Standard Model of particle physics, which dictates that the mass eigenstates are themselves $|\nu_e\rangle, |\nu_\mu\rangle$, and $|\nu_\tau\rangle$. This is one of two experimentally observed neutrino phenomena that are beyond-the-Standard-Model (BSM), the other being the measurement of their nonzero mass.

Neutrino oscillations occur over macroscopic distances and so are extremely unlikely to occur over the length of the detectors considered in this thesis.

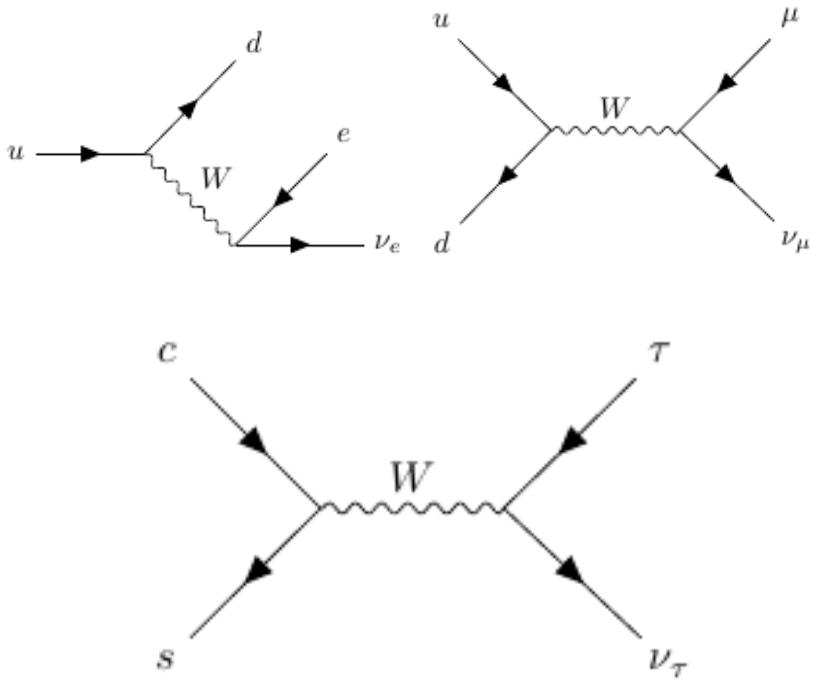

Figure 2.3: Feynman diagrams for dominant production mechanisms for ν_e, ν_μ, and ν_τ.

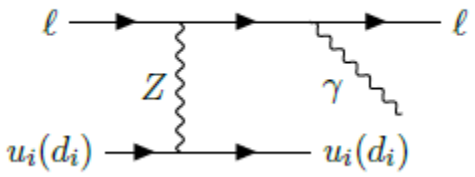

Figure 2.4: Feynman diagrams for electron bremsstrahlung.

2.3 Electron interactions

The lepton interactions will be important because they are the means by which the flavour of the interacting neutrino will be identified in FASERν. Among the charged leptons, the

electron is most important because, upon contact with nuclear matter, high-energy electrons will lose much of their energy through *bremsstrahlung,* or braking radiation, in the form of photons. These photons will interact with matter to stimulate production of electron-positron pairs, each of which will also undergo bremsstrahlung provided there is sufficient energy. This process is known as an *electromagnetic shower,* and continues until the energy falls below the *critical energy*, E_c, of the absorber; at that point, the dominant mechanism for loss of the electron energy becomes the collisions with the nucleons themselves.

2.4 Muon interactions

The muon interactions will be one of the primary sources of background. Muons will be produced at the LHC and travel through the rock. The muon is approximately 207 times heavier than the electron, and can decay via the process $\mu \rightarrow e\bar{\nu}_e\nu_\mu$; however, this process takes 2.197×10^{-6} s [9], during which the muon travels an extended distance. Therfore, for the studies addressed in this thesis, the muon will be treated as a stable particle.

Muon interactions have been studied at low energies [13] as well as high energies [14]. As with electrons, muons are also susceptible to bremsstrahlung. However, due to their increased mass, they will radiate much less energetic photons and retain enough momentum afterward to do so repeatedly. These low-energy photons are unlikely to produce electron-positron pairs upon further interactions with matter. Therefore it is clear that muons which induce high-energy electrons will not do so through bremsstrahlung and pair production.

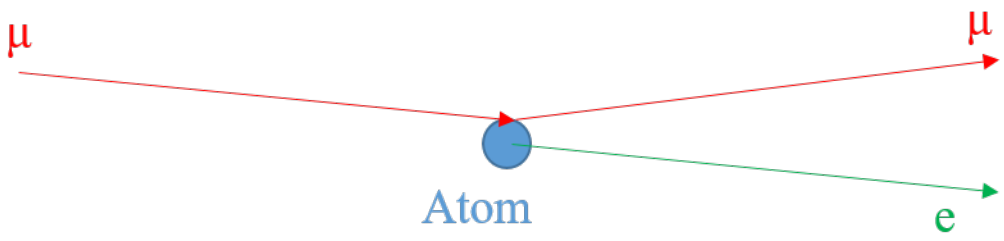

Figure 2.5: Schematic illustration of a muon-induced knock-on electron production event.

At higher energies, muon interactions with matter are dominated by the production of *knock-on electrons*. The process is illustrated diagrammatically in Figure 2.5: A muon (red) interacts with an atom (blue), causing the muon to be deflected and the atom to release one of its electrons (green). At energies above $E \approx 1\,\text{TeV}$, this process is the preferred mechanism by which muon interactions with matter produce electrons over bremsstrahlung.

Unlike the electron and the muon, the τ particle decays in $2.903 \pm 0.005) \times 10^{-13}$ s, which is a short enough time and corresponds to a short enough distance that τ particles produced in FASERν will likely decay within FASERν, usually to a lighter lepton or charged pion. One

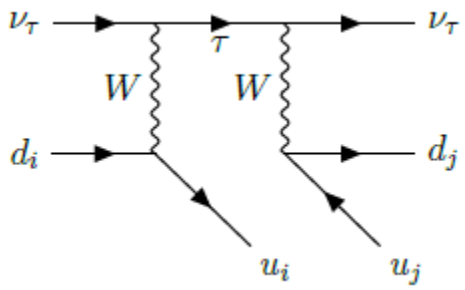

Figure 2.6: Feynman diagram for a ν_τ CC event producing a τ which promptly decays to a pion.

particular neutrino CC interaction that is worthy of note is the ν_τ CC event, which produces a short-lived τ particle that subsequently decays back to a ν_τ and a π^- meson. The Feynman diagram for this process is shown in Figure 2.6. As will be discussed subsequently, this is the signature of tau neutrino CC events which will allow it to be distinguished from the neutrino CC events of other flavours. It is also possible for the tau to decay to a muon and two neutrinos, which will be relevant for charge identification.

Although the interactions of neutrinos is important for the later study of neutrino-nucleon cross-section sensitivity, it will not be discussed in this thesis, which is concerned with electron clustering and energy regression as well as tracker performance (which are both prerequisites for studying neutrinos with the FASER detector).

Chapter 3

EXPERIMENTAL APPARATUS

3.1 The Large Hadron Collider

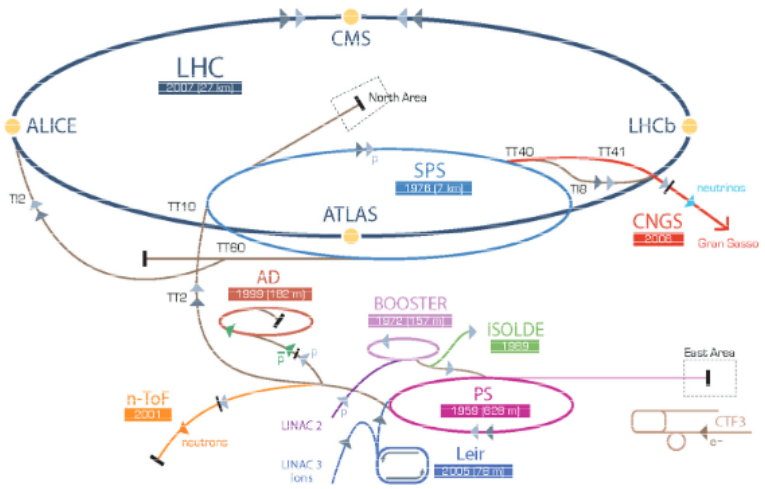

Figure 3.1: Schematic diagram of the Large Hadron Collider, as well as the four major detectors ATLAS, CMS, ALICE, and LHCb positioned along the collider ring and the infrastructure to produce the high-energy proton beams.

The Large Hadron Collider (LHC) is a 27-km particle accelerator on the French-Swiss border. Affiliated with the European Center for Nuclear Research (CERN), it collides proton beams in opposite directions. The LHC was upgraded to collide protons at center-of-mass energy $\sqrt{s} = 13$ TeV in April 2015, an upgrade from $\sqrt{s} = 8$ TeV previously.

The LHC infrastructure is described in Figure 3.1 [15]. Hydrogen atoms are separated into their constituent protons and neutrons with an intense electric field. Then an alternating field is imposed with a specified frequency so that the charged particles are directed along the

LHC, with the frequency controlled to allow protons to arrive in so-called "bunches" rather than as a continuous stream. These bunches have only a modest, nonrelativistic kinetic energy, so they are then injected into a linear accelerator and subsequent proton synchrotron booster (annotated in the figure as LINAC and BOOSTER, respectively) which accelerates the protons and antiprotons up to 50 MeV. They then enter the proton synchrotron (abbreviated as PS above) which further boosts the protons to 25 GeV. Some antiprotons are diverted to the antiproton decelerator (AD) for use in antiproton experiments; the rest are sent with the accompanying protons to the super proton synchrotron (SPS) which boosts them to 450 GeV. The protons and antiprotons are then passed to the main collider ring of the LHC, where they are accelerated to their final energies of 6.5 TeV.

In order to accelerate the individual protons to energies of 6.5 TeV, there are approximately $10,000$ superconducting copper-clad niobium-titanium magnets weighing ~ 470 metric tonnes [16], with some 27 metric tonnes in the $1,232$ dipole magnets keeping the particles contained within the collision ring. Quadrupole magnets are used as corrections to the dipole magnetic field in order to maintain beam focus so as to maximize the probability of the proton beams crossing and interacting. Octopole and higher-order multipole magnets are used as further corrections to the magnetic field [15].

The protons travel in two parallel beam pipes. When they have attained sufficient energy, the beams are steered to collide and interact with one another at the nominal center-of-mass energy. This can be made to happen at one of four interaction points. These interaction points are calibrated to occur inside the detector volumes of the four main experiments positioned around the collider ring: ATLAS, CMS, ALICE, and LHCb [17]. In addition to these experiments, Protons from the LHC are also used in an array of beam-dump experiments [18] and fixed-target experiments, such as the proposed SHiP experiment [19].

3.2 The FASER Spectrometer

The ForwArd Search ExpeRiment, or FASER, is designed to detect new weakly-interacting particles with low mass. One such candidate particle is the dark photon decaying to photons,

e^+/e^-, or $\mu 6+/\mu^-$ [20]. Given the anticipated low mass, any such particles produced at the ATLAS interaction point (IP) will continue on a nearly tangent path to the LHC collider ring. There is a site in the TI12 maintenance tunnel which intersects the line-of-sight to the ATLAS IP; this is the site where FASER is presently being installed. A trigger and data acquisition (TDAQ) system is employed to ensure events of interest are recorded.[21] Figure 3.2 illustrates the region downstream from ATLAS as well as the LHC infrastructure

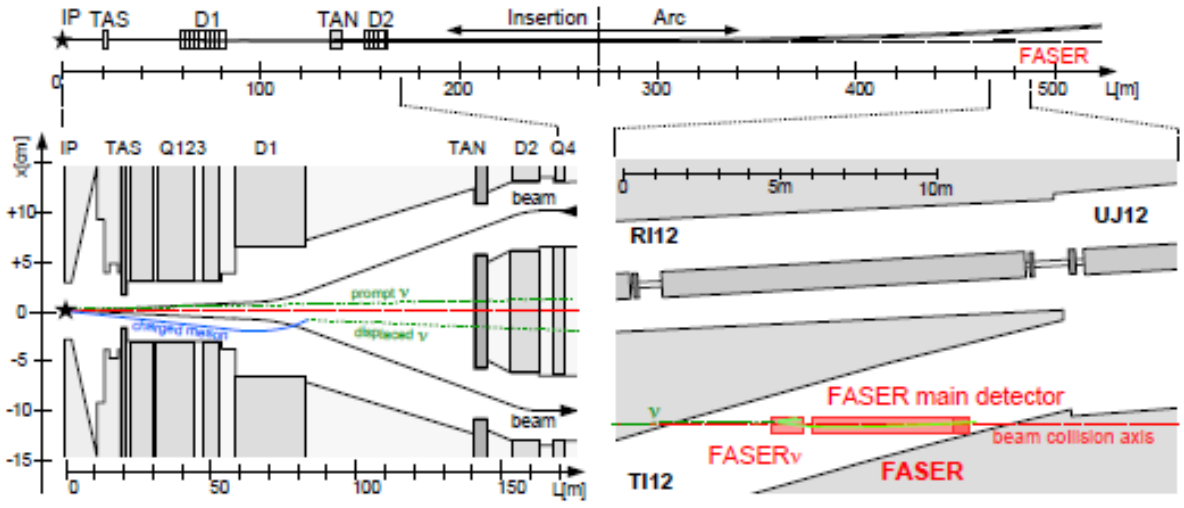

Figure 3.2: Schematic diagram of the region downstream of the ATLAS IP. The vertical axis is scaled disproportionately for clarity. Charged particles are deflected by the LHC magnets while neutral particles continue on a straight line until they are absorbed by the TAN neutral particle absorber. Neutrinos, dark photons, and other new physics particles may continue uninterrupted, travelling approximately 480 m until they reach FASER.

along the trajectory of the particles emerging from the ATLAS IP.[2] After the proton-proton collisions at the ATLAS IP, many particles are produced, including charged particles, neutral hadrons, neutrinos, and new physics particles. The charged hadrons and leptons are deflected by the LHC quadrupole magnets indicated. The neutral hadrons travel further, but are absorbed by the TAN neutral particle absorber. This leaves the neutrinos, dark photons, and other new physics particles. These may travel through the TAN without any interaction,

and therefore are likely to continue travelling through the 480 m of rock remaining before reaching the TI12 maintenance tunnel and the FASER spectrometer. The positioning of FASER relative to the collider ring is illustrated in Figure 3.3

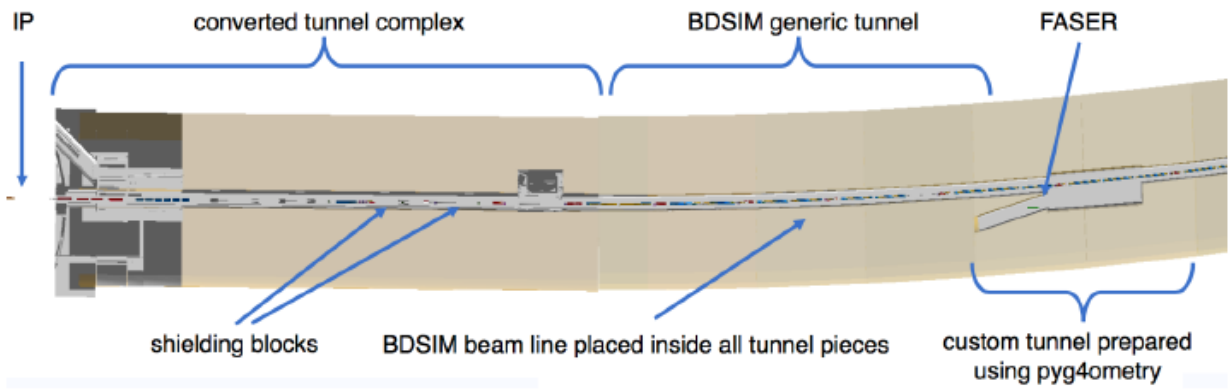

Figure 3.3: Schematic diagram illustrating the placement and orientation of FASER relative to the collider ring of the LHC.

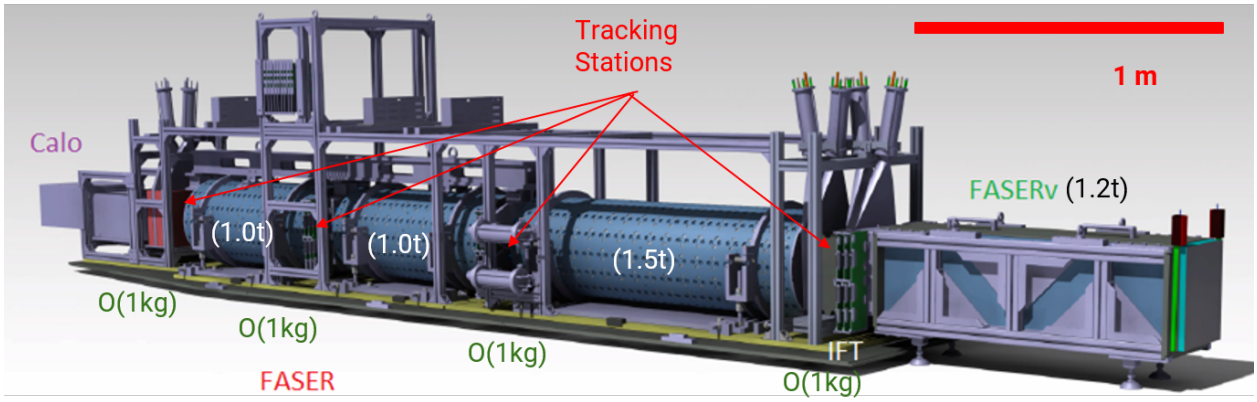

Figure 3.4: Computer-generated model of the FASER spectrometer (red), FASERν emulsion detector (green), IFT (white), and calorimeter (purple). The tracking stations are indicated with red arrows. The masses of each detector element are displayed in parentheses.

Figure 3.4 shows a computer-generated schematic of the FASER spectrometer.[22, 23] The detector itself consists of a veto scintillator, three tracking stations, and a hybrid calorimeter.[24] The veto scintillator serves the purpose of eliminating 99.99 percent of muons incident on FASER. The calorimeter is a hybrid hadronic and electromagnetic calorimeter that will detect hadrons and photons.

Figure 3.5: FASER spectrometer installed in the TI12 maintenance tunnel of the LHC for commissioning as of 5 February 2021.

Figure 3.5 shows the real FASER detector as it has been installed in the TI12 maintenance tunnel of the LHC. The FASER spectrometer is currently in the process of commissioning.[2]

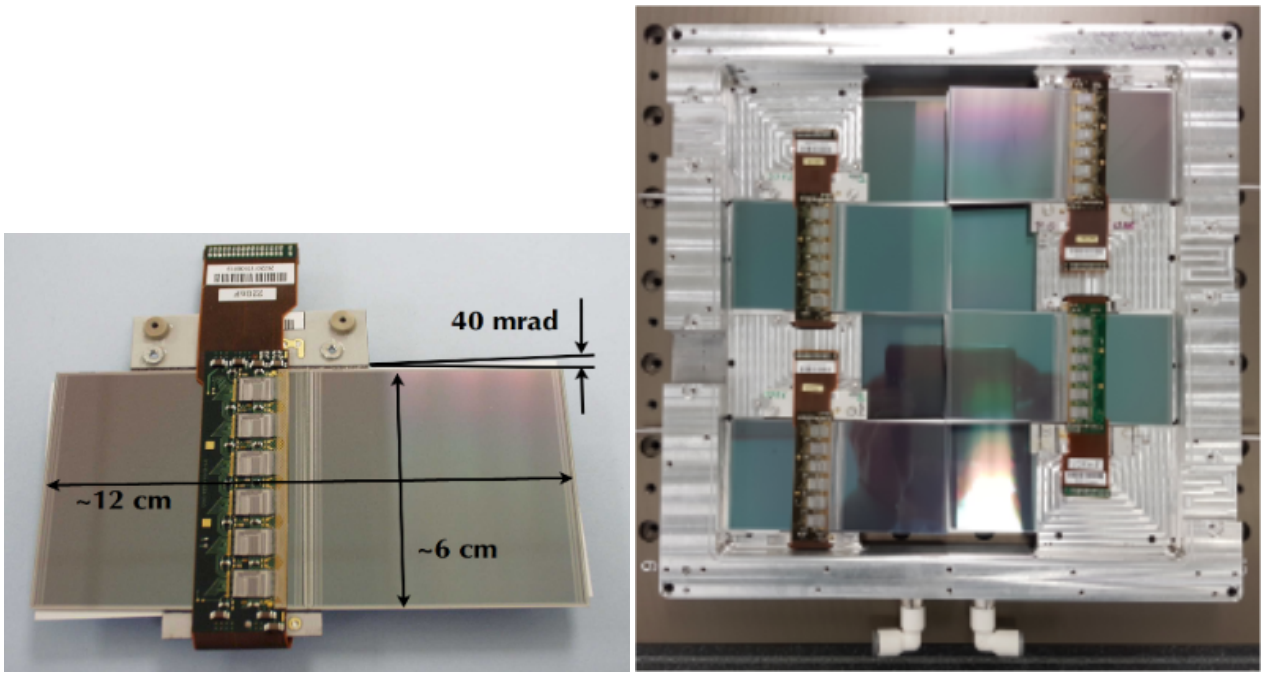

Figure 3.6: (left) SCT module; (right) Eight SCT modules assembled into one FASER tracking plane.

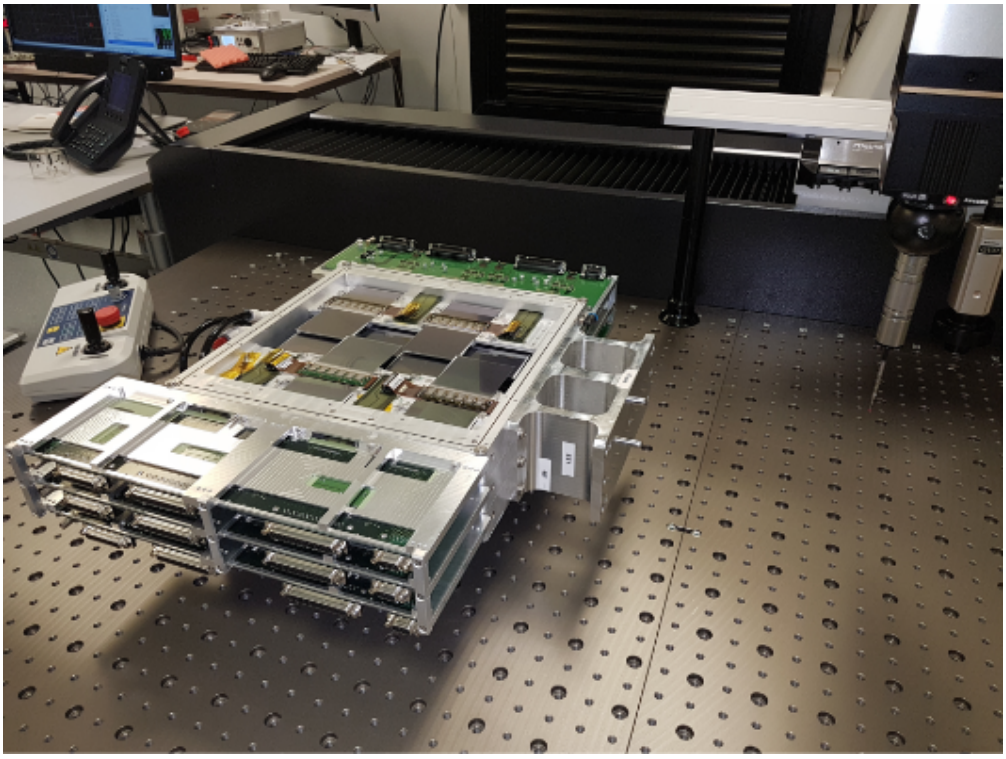

Figure 3.7: One of the three SCT tracking stations assembled from three tracking planes.

The tracking stations each contain three tracking planes. Each tracking plane consists of eight staggered semiconductor tracker (SCT) modules [25], as illustrated in Figure 3.6.[2] Three tracker planes are assembled into an SCT tracking station, as in Figure 3.7[2] where a tracking station is undergoing metrology to determine the exact dimensions (for input into the simulation).

. Each SCT module consists of 128 strips on each side of the module which record hits from particles passing through them. The SCT strips on opposite sides of the same module are tilted relative to one another by a stereo angle of 40 mrad, which allows the combination of strips on opposite sides of the SCT module to resolve the position of the recorded hit.

To aid in the separation of charged particles, there is a 0.55 T magnetic field which is sourced by the dipole magnets in the volumes between the tracking stations. As a result of this magnetic field, charged particles traveling through FASER will leave linear tracks in the tracking stations, but helical tracks in the dipole volumes.

The FASER detector will also include four scintillator stations. Each scintillator station consists of two identical 20 cm $\times$ 20 cm $\times$ 2 cm scintillator layers connected to a photomultiplier tube (PMT) through a light guide. The detection efficiency for a single scintillator layer is well in excess of 99%.

The purpose of these two veto stations is primarily to eliminate events with incoming particles, namely high-energy muons. These muons may undergo muon bremsstrahlung upstream from the FASER spectrometer producing photons which go undetected.

To resolve this, two veto scintillator stations will be placed upstream from FASER. The scintillator stations are separated by a block of lead which acts as an absorber. This leads to the following behaviour:

- A photon passing through the first veto scintillator station will be entirely contained within the lead absorber block.

- The photon will initiate an electromagnetic shower, producing $e^+/e^-/\gamma$, and the photons from this shower will be detected by the second veto scintillator station.

In either event, the veto scintillator stations are sufficient to eliminate over 99.99% of incoming muons from entering the FASER spectrometer.

In addition to the SCT tracking stations, the design for the FASER detector also includes an electromagnetic calorimeter. This is a detector component whose purpose is to measure the deposited electromagnetic energy – i.e., from electrons and photons – without making a clear distinction of individual particle energies.

3.3 The Interface Tracker

The interface tracker (IFT) is a fourth tracking station that will be placed upstream of the FASER spectrometer but downstream of the FASERν emulsion detector, as shown in Figure 3.4 above. As the name implies, the IFT interfaces FASER to FASERν as a fourth tracking station. The IFT will improve the tracking performance of the FASER spectrometer. In particular, a better position resolution is expected with FASER+IFT than with FASER alone. There is also better charge reconstruction resolution, as more points are available to be used in the reconstruction of the helix track.

3.4 The FASERν Emulsion Detector

FASERν is a prospective emulsion detector that will be placed upstream from FASER and is designed to detect neutrinos. Neutrinos have been detected from beam-dump experiments such as NA62, as have cosmic neutrinos by experiments such as IceCube in Antarctica, but importantly no neutrino produced at a collider has ever been detected.

The emulsion detector will consist of 760 layers of tungsten absorber. In between consecutive absorber layers are emulsion films on either side of a plastic base. The emulsion films use silver bromide crystals that will record hits from any charged particles. These films will be removed and developed for analysis during Run 3 of the LHC [26]. Figure 3.8 shows the anatomy of one of the emulsion films, as well as a track left by a minimum ionizing particle from a 10 GeV pion beam.[2] Figure 3.9 shows the support structure that will hold the 760 alternating emulsion films and tungsten absorbers of FASERν.[2]

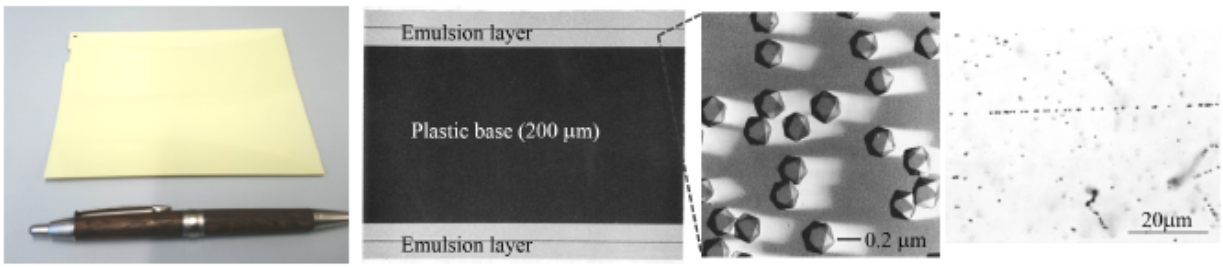

Figure 3.8: (far-left) Rectangular section of an emulsion film; (center-left) schematic diagram of emulsion film showing two emulsion layers on either side of a 200 μm plastic base; (center-right) electron microscope photograph of AgBr crystals in emulsion film; (far-right) image of the emulsion track left by a secondary particle emerging from a 10 GeV π.

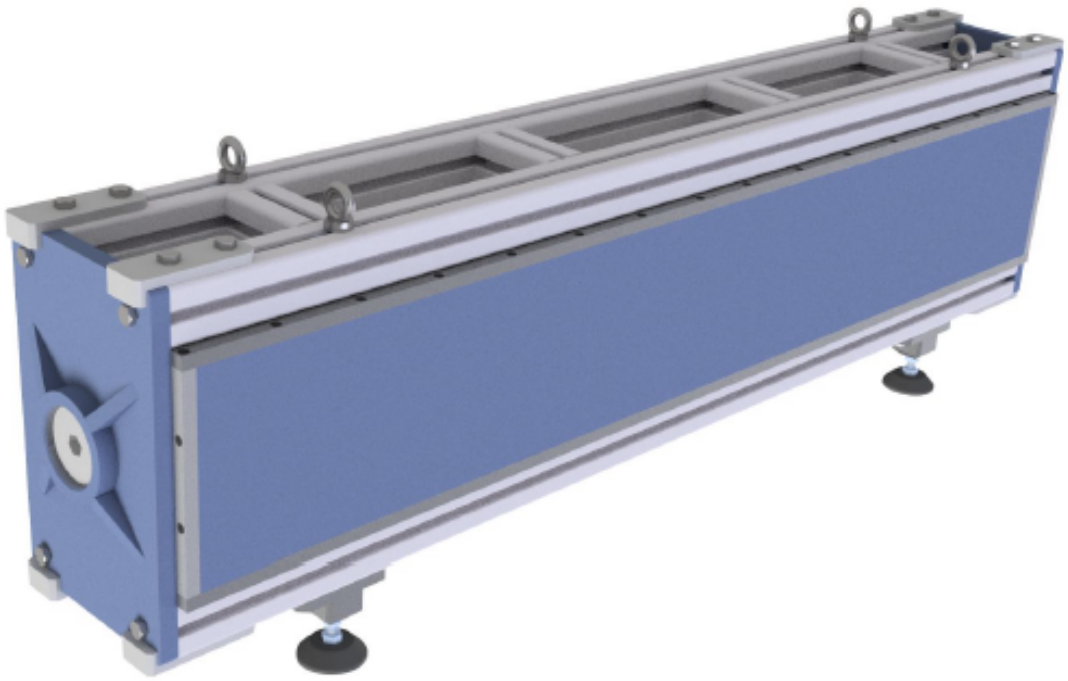

Figure 3.9: Schematic diagram of support structure that will contain the FASERν emulsion detector.

FASERν will be able to distinguish between ν_e, ν_μ, and ν_τ CC events. Since neutrinos will not interact with the emulsion detector, but the secondary charged leptons will interact in different ways according to their flavours, these events can be distinguished using the following characteristics:

- An electron passing through an absorber will undergo *bremsstrahlung*, giving off an energetic photon which can split into an $e+/e-$ pair upon interacting with the absorber. This process continues until the dominant mechanism of energy loss becomes the collisions. This is known as an *electromagnetic (EM) shower.*

- A muon passing through an absorber will similarly undergo bremsstrahlung, but due to the muon's mass the photons will be much less energetic, and as a result the EM showers will be much smaller.

- A tau passing through an absorber will rapidly decay, producing another ν_τ and either a μ, ν_μ pair or a charged pion.

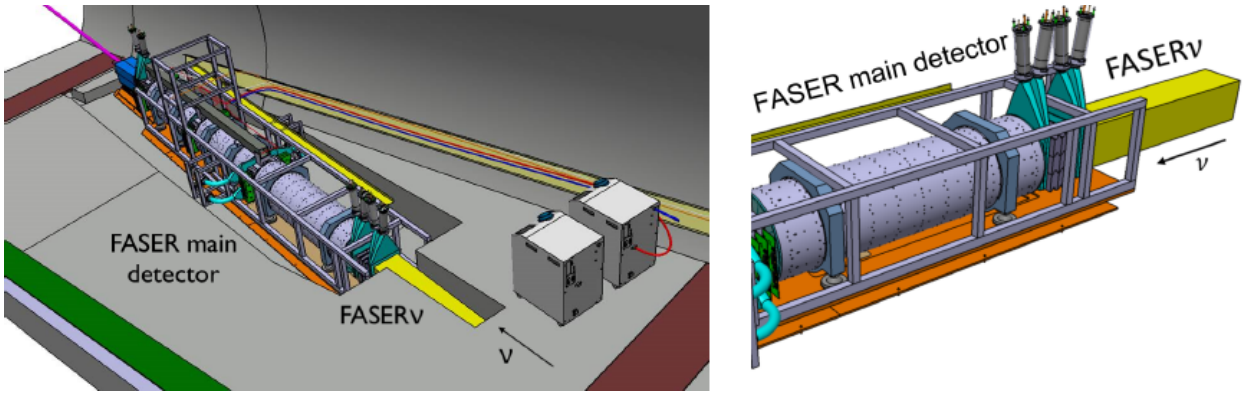

Figure 3.10: Rendering of the placement of the FASER spectrometer and the FASERν emulsion detector in the TI12 tunnel. The ATLAS IP is off-screen to the right.

Figure shows the placement of the FASER spectrometer and the FASERν emulsion detector in the TI12 maintenance tunnel, with the ATLAS interaction point off-screen to the

Table 3.1: Estimated number of neutrino/antineutrino CC interactions in FASERν for each of the three flavours of neutrino. [27]

	e	μ	τ
ν	2986	8437	110
$\bar{\nu}$	1261	2737	55

right.

Although the FASERν detector will be able to distinguish between the three different flavours of neutrino interactions, it will be unable to distinguish neutrino CC events from antineutrino events. This is because the signatures of ν_e, ν_μ, and ν_τ events in terms of emulsion tracks are precisely the same as their charge-conjugated counterparts. Therefore, what will actually be measured is not the separate neutrino- or antineutrino-nucleon cross-section, but rather the weighted sum $N_\nu\sigma(\nu N)+N_{\bar{\nu}}\sigma(\bar{\nu}N)$, where N_ν and $N_{\bar{\nu}}$ are the expected numbers of neutrino and antineutrino events for the given flavour, respectively.

Because the IFT is immediately downstream from FASERν, it is now possible to identify tracks left in the FASER spectrometer with tracks left in the emulsion detector. This process is referred to as *track matching*. Without the IFT, the extended distances the particles travel before recording hits in FASER makes it impossible to match FASER and FASERν tracks. This functionality will prove essential to the separate measurement of neutrino and antineutrino cross section measurements, as will be described later.

For FASERν, the number estimated number of expected neutrino and antineutrino CC interactions that will occur in the FASERν emulsion detector during Run 3 of the lHC is shown in Table 3.1. These estimates were obtained using the `DPMJET` event generator [28], and are a revision of those presented in [26] which were obtained using the `EPOS-LHC` [29], `QGSJET` [30], `SIBYLL` [31], and `Pythia8` [32] generators.

The FASERν emulsion detector will offer unparalleled position resolution. Unlike the FASER spectrometer, which records hits in individual silicon strips of a single SCT module,

the emulsion detector uses an emulsion of silver bromide (AgBr) crystals with mean radius $200\,\mu$m, as shown in the right-most panel of Figure 3.8. Each crystal independently functions as a detection channel for charged particles, and there is an extremely high density of these channels ($\sim 10^{14}\,\mathrm{cm}^{-3}$). These crystals are arranged into emulsion layers approximately $50\,\mu$m thick, on opposite sides of a $200\,\mu$m-thick plastic base to form an emulsion film. The FASERν detector will use 760 emulsion films alternating between absorber layers 1 mm thick.

The choice of absorber material is specified to be a high-density material to maximize the number of nucleons in the FASERν detector volume; this is because, given the interaction cross section σ, the probability of an interaction is

$$P = \sigma \int L\,\mathrm{d}t, \tag{3.1}$$

where $\int L\,\mathrm{d}t$ represents the integrated luminosity and has dimensions of inverse area. The number of expected interactions in FASERν is proportional to the number of scattering centers (nucleons), which means that the absorber material should be chosen to satisfy two criteria:

- The material should have as high an atomic mass N as possible to maximize the number of nucleons per target nucleus.

- The material should have as high a density ρ as possible to maximize the number of scattering centers in a target volume.

Based on these two criteria, two absorber material candidates emerge, namely tungsten ($N = 184$, $\rho = 19.3\,\mathrm{g/cm}^3$) and lead ($N = 208$, $\rho = 11.35\,\mathrm{g/cm}^3$). There is no timing information available for the emulsion detector; hits recorded will simply continue to pile up in the detector until the emulsion films are replaced. This process involves transporting the entire emulsion detector to the surface, disassembling it, and processing the emulsion films. The replacement films are prepared in advance so that they may be reassembled into the FASERν detector quickly.

The emulsion reconstruction technique works for track densities up to approximately 10^6 tracks per square centimeter. As such, the emulsion films need to be replaced once for every $20 - 30\,\mathrm{fb}^{-1}$ of data collected. Given the luminosity of the LHC during 2021 and the expected luminosity for 2022-2023, seven sets of emulsion films were prepared, with one set to be implemented in 2021 and three each in 2022 and 2023.

3.5 The Pilot Detector

The FASERν pilot detector was installed in the LHC from September 2018 to October 2018. It consists of two modules; the lead module consists of 101 layers of lead interleaved with emulsion films, whereas the tungsten module consists of 122 layers of tungsten interleaved with emulsion films. Figure 3.11 shows the lead and tungsten modules, labeled as Pb and

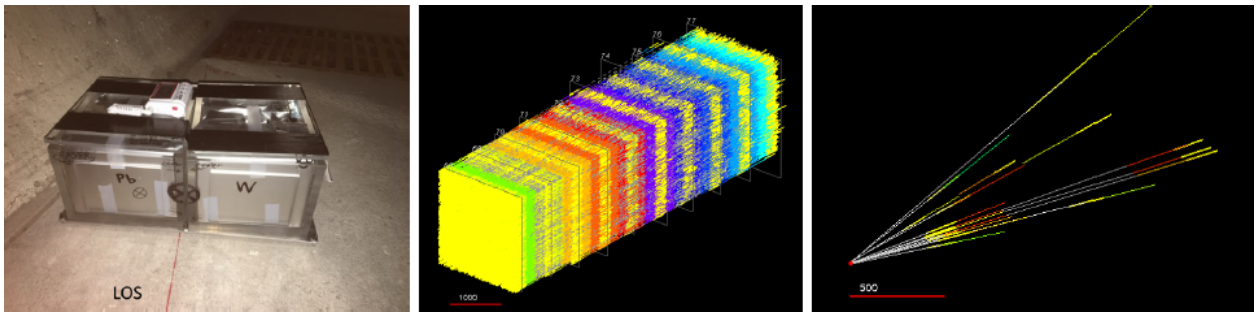

Figure 3.11: The two tungsten (W) and lead (Pb) modules of the FASERν pilot detector placed in TI12 in September-October 2018.

W, respectively.[2] The modules are placed so that the line-of-sight to the ATLAS IP is contained in the plane separating the modules (marked by an X at the border of the two modules). This pilot detector collected $12.5\,\mathrm{fb}^{-1}$ of data during this run, during which a track density of 10^6 tracks/cm^2 was measured, as shown in the middle plot. On the right is a reconstructed candidate neutrino event.

Unlike the FASERν emulsion detector, there is no lepton-flavour identifying capability with the pilot detector. This is because the lepton ID is done in the last 100 layers of

FASERν, so there is insufficient detector hardware to perform lepton-flavour identification. Nevertheless, as indicated on the right-most panel of Figure 3.11, there was a candidate neutrino event identified during the pilot run. The analysis of the FASERν pilot run detector will be discussed in a later chapter.

Chapter 4

TRACKER AND EMULSION DATASETS

4.1 Tracker Performance Dataset

To evaluate the tracking performance of the FASER spectrometer, it is necessary to prepare MC samples of muons, which will register hits in the detector that can be used to reconstruct their tracks. This is accomplished using the Calypso software framework in conjunction with ATLAS and Athena. This official FASER simulation software may be downloaded, installed, and run on `LXPLUS`.

For this study a sample of $5 \times 10^5\,\mu^+$ and $5 \times 10^5\,\mu^-$ was prepared, with uniform energy distributions from 0 to 5 TeV and angular distribution (with respect to the axis of the detector from 0 to $\pi/20$. (The samples are prepared as 100 5×10^3 muon samples due to memory usage limitations of the software.) These samples are then run through the full chain of digitization and combinatorial Kalman filtering to produce the truth and reconstructed tracks used for the performance study.

4.2 Pilot Data Sample

For the pilot run in 2018, $12.2\,\mathrm{fb}^{-1}$ of data were collected with the two FASERnu pilot detector modules. This pilot run has recorded the first neutrino interaction candidates at the LHC, as reported in [33]. There is a background of muons incident on the pilot detector at a density of approximately 2.1×10^5 muons per square centimeter, with the charge distribution being roughly even between positive and negative muons.

For each basetrack recorded in the FASERnu pilot detector, there are two angles of interest, namely $\theta_x = \arctan(p_x/p_z)$ and $\theta_y = \arctan(p_y/p_z)$. What is recorded by the emulsion detector is not θ_x and θ_y, but rather $Tx = \tan\theta_x$ and $Ty = \tan\theta_y$. These are

the slopes the basetrack makes with the z-axis in the (z, x)- and (z, y)-planes, respectively. It is clear that the angular distributions of basetracks should peak at $Tx = Ty = 0$, since the basetracks will be recorded by particles originating from primary particles produced at the ATLAS interaction point 470 meters away, and the FASERnu pilot detector is only $12.5\,\text{cm} \times 10\,\text{cm}$ in area.

The study of interest with the FASERnu pilot run is that of electron energy reconstruction in the pilot detector. The reasons for this are twofold:

- The energy spectrum of positive and negative muons incident on the FASERnu pilot detector from the ATLAS interaction point has been obtained via dedicated simulations. However, the actual data from the FASERnu pilot run allows for the validation of the simulation spectrum. The high-energy muons incident on the pilot detector will occasionally interact with atoms to produce a secondary electron (so-called knock-on electrons or δ-rays). By simulating the expected muon flux incident on the FASERnu pilot detector, it is possible to identify and reconstruct the EM showers of these knock-on electrons and therefore their energies. By applying the same procedure to the pilot run data and pilot run MC simulation, the number of identified knock-on electrons can be compared between data and simulation, and the energy spectra compared. If the knock-on electron energy spectra are consistent, the muon energy spectrum will be validated.

- In order to reconstruct the energy of an electron neutrino CC event, the energy of the outgoing electron must be reconstructed. It is known that higher energy electrons will leave larger clusters (that is, clusters containing more e^+/e^- basetracks), and in particular the number of truth e^+/e^- basetracks in an electron EM shower grows approximately linearly with energy. A well-designed electron clustering algorithm should produce a similar linear relationship between clustered basetracks and truth electron energy. Using this, we can reconstruct the electron energy and thus the neutrino energy by counting the number of basetracks associated to the cluster, and therefore measure

the electron neutrino CC cross section by counting the number of electron neutrino interactions in FASERnu in each energy bin.

The data recorded by the FASERν pilot detector is stored in the FEDRA format, containing the following information:

- Position (x, y, z), angular $(\tan\theta_x, \tan\theta_y)$, and quality-of-fit for microtracks used for emulsion reconstruction.

- Position, angular, and quality-of-fit for basetracks reconstructed from the above microtracks.

These data are stored in 30 subvolumes for the pilot detector, where each subvolume is a square cylinder with cross-section $22\,\text{mm} \times 22\,\text{mm}$ and overlaps with neighbouring subvolumes by 2 mm. For the electron clustering study, the last 29 (highest-quality) plates are used.

4.3 Pilot Detector MC Simulation Sample

In order to develop an electron clustering algorithm that can be used on the FASERν pilot data, an appropriate simulation data set must first be studied. For maximum consistency with the pilot data, this sample must use the correct track density as well as the simulated muon energy and angular spectra.

To determine the correct energy and angular spectra, a GEANT4 simulation of the rock immediately upstream from the pilot detector was prepared. The input energy and angular spectra to this rock simulation were those already simulated originating from the ATLAS IP and traveling 409 m to the location in question. The energy spectrum of muons arriving at the pilot detector is shown in Figure 4.1. Note that there is a peak in the energy spectrum of μ^- around $\sim 1\,\text{TeV}$ that is not present in the μ^+ spectrum. The angular distributions for μ^+ and μ^- at the pilot detector are shown in Figure 4.2.

The GEANT4 simulation was carried out with 2×10^7 muons, approximately the expected density of muons arriving at the FASERν pilot detector. The first step of the analysis was

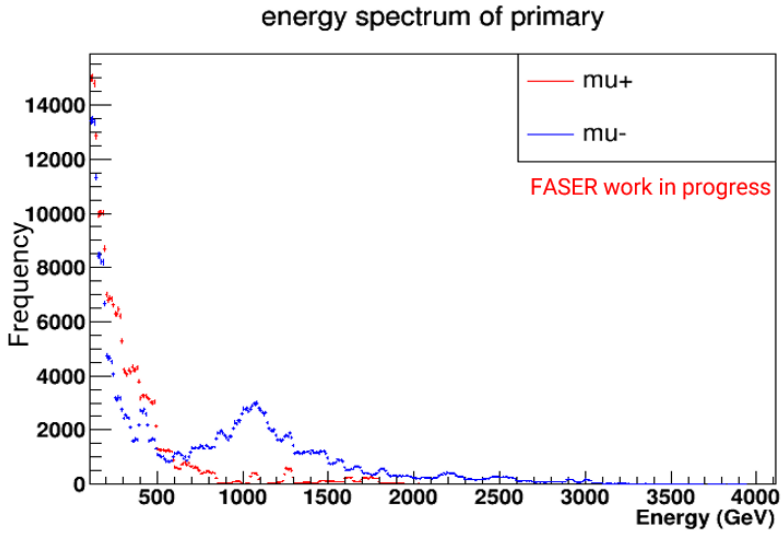

Figure 4.1: Energy spectra of μ^+ (red) and μ^- (blue) arriving at the FASERν pilot detector.

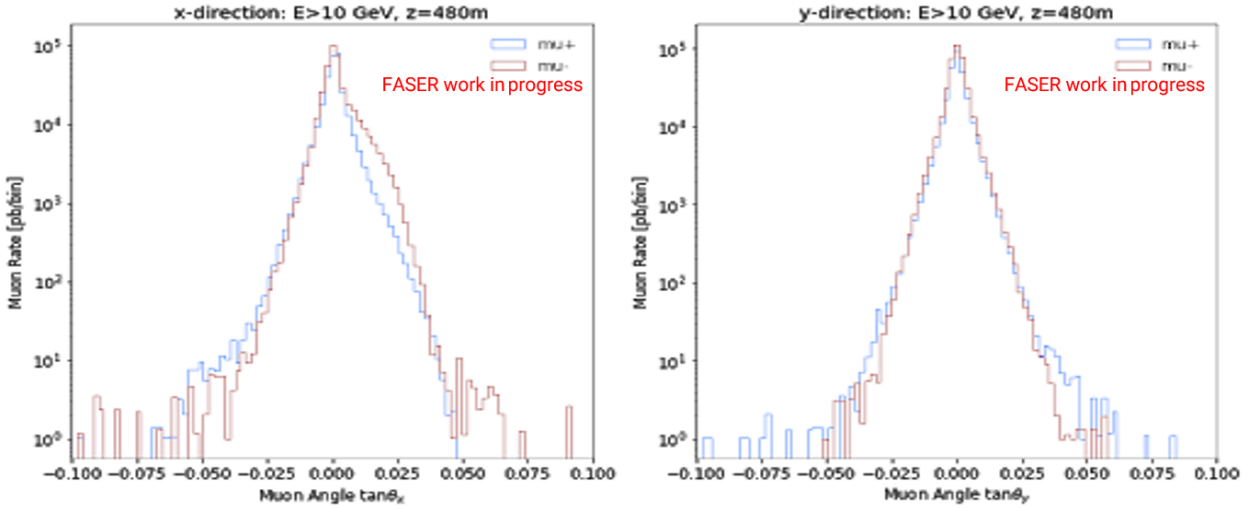

Figure 4.2: Spectra of $\tan\theta_x$ and $\tan\theta_y$ for μ^+ (red) and μ^- (blue) arriving at the FASERν pilot detector.

to preprocess the simulation data to remove all unnecessary information. The G4 hits were retained with position (x, y, z) and angular $(\tan\theta_x, \tan\theta_y)$ information, as well as the truth energy of the particle and the truth event index to which it belongs. (In the following

chapter, the pilot run data will have been prepared in the same format, without the truth-linked information, so that the clustering algorithm may be run on both data sets.)

In order to develop an electron clustering algorithm that can be used on the FASERν pilot data, an appropriate simulation data set must first be studied. For maximum consistency with the pilot data, this sample must use the correct track density as well as the simulated muon energy and angular spectra.

The samples produced by the GEANT4 simulation are very large, memory-intensive files ($\sim 1\,\mathrm{TB}$) and must be reduced in size and complexity before they can be studied. At truth level, the GEANT4 simulation stores a large amount of data either not needed for the electron clustering study or not available in the pilot data. The only data which are used for the clustering algorithm are

1. Position of the midpoint of the basetrack (x, y, z).

2. Slopes of the basetrack in the $(x, z)-$ and $(y, z)-$planes (Tx, Ty).

3. Truth energy and PDGID of the particle recording the basetrack.

4. Truth event containing the basetrack.

In the above, (1) and (2) are used for the electron clustering study, while (3) is for the electron energy regression study and (4) is for electron identification.

Chapter 5

TRACKING AND RECONSTRUCTION WITH FASER

This chapter addresses the mechanics of tracking and reconstruction with the FASER spectrometer and the FASERν emulsion detector. The workflow for tracker reconstruction in FASER and emulsion reconstruction FASERν is described and the performance of the tracker reconstruction algorithm is presented. The legacy `FaserMC` GEANT4 simulation is briefly introduced and the results of the legacy track fitting and new hybrid helix/linear track fitting are presented. The current tracking model implemented in calypso is discussed. The digitization format used by the SCT modules is outlined. The clustering algorithm is presented and the use of pairs of clusters on opposite sides of SCT modules for space-point formation is discussed. Finally, the track-finding and track-fitting algorithms and Kalman filters are discussed.

5.1 Tracking in FASER

Particles passing thorugh FASER will register hits in the SCT modules which comprise each of the four tracking stations (including the iFT). The task of *tracker reconstruction* is to use these hits to reconstruct the tracks.

5.1.1 Digitization

The first task is to convert the hit information recorded in the SCT modules to an efficient digitized format. Due to the overwhelming density of particles expected to be incident on FASER from the ATLAS IP, the detailed hit information that is output by the SCT module is in a bitstream format which maximizes the efficiency of bits used for data [34]. The bitstream format is as shown above in Figure 5.1. The header and trailer are key bit sequences which

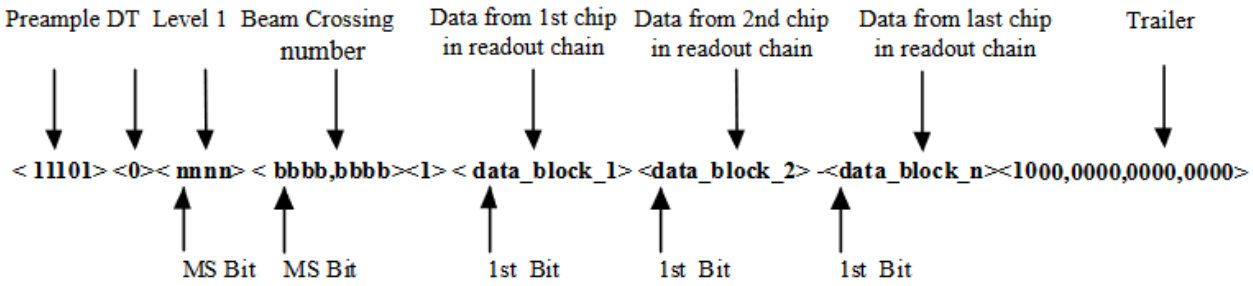

Figure 5.1: Bitstream format for readout from the SCT modules in the FASER spectrometer.

contain between them a series of data blocks that specify how many strips in the SCT module registered hits. There are three distinct patterns depending on if there is a hit in an isolated strip, multiple adjacent strips, or no strips at all in the SCT module.

5.1.2 Clustering

Because the bitstream encodes adjacent strips that have been hit, it is possible to combine them into "hit clusters." Since each SCT module has 128 strips on each side, the cluster information stores information about a particle's position in one of the two transverse directions. This is illustrated schematically in Figure 5.2. The clusterization algorithm must take into account so-called "bad strips," which are strips that are damaged or otherwise malfunctioning, before combining all hits in adjacent strips into a cluster.

5.1.3 Space-Point Formation

Using the cluster information reconstructed in the previous step, the position of a particle can be resolved in one transverse direction only. However, there are strips on both sides of the SCT module, and these are tilted at a stereo angle of $\delta\theta = 40$ mrad, as indicated in Figure 5.3.Therefore, the cluster information recorded on both sides of the SCT module provide position information in two linearly independent directions and, by combining them, it is possible to reconstruct a point in space through which the particle registering the hits

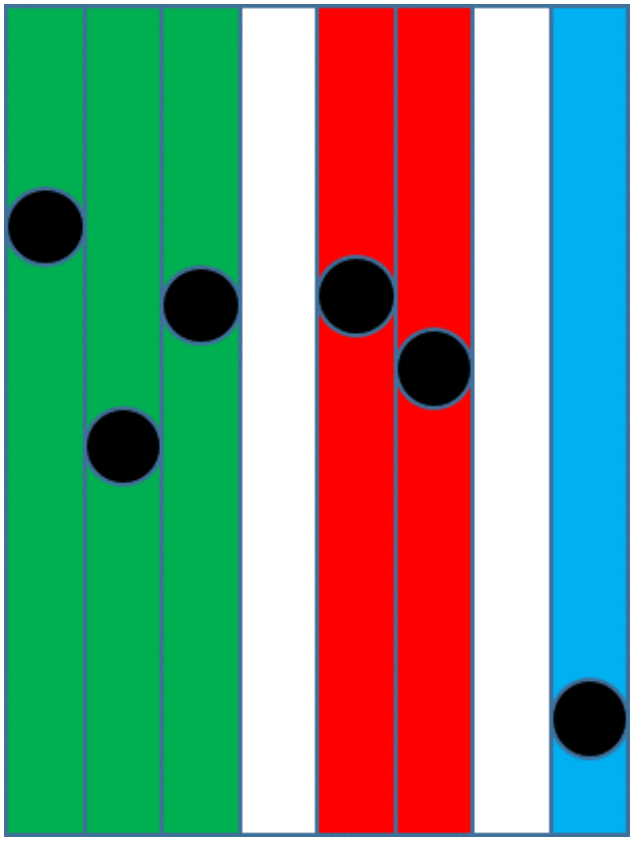

Figure 5.2: Schematic diagram of hits being combined into hit clusters on one side of a given SCT module. The 3 green hit strips are combined into a single hit cluster, as are the two red hit strips and the single blue hit strip.

passed. This point is known as a *space-point* and is an essential ingredient in the tracker reconstruction.

It is important to note that, because of the small stereo angle, the position resolution of the particle is substantially better in one direction than in the other. Resolution in the direction perpendicular to the SCT strips is $\sim 16\,\mu$m, whereas resolution in the direction parallel to the SCT strips is $580\,\mu$m.

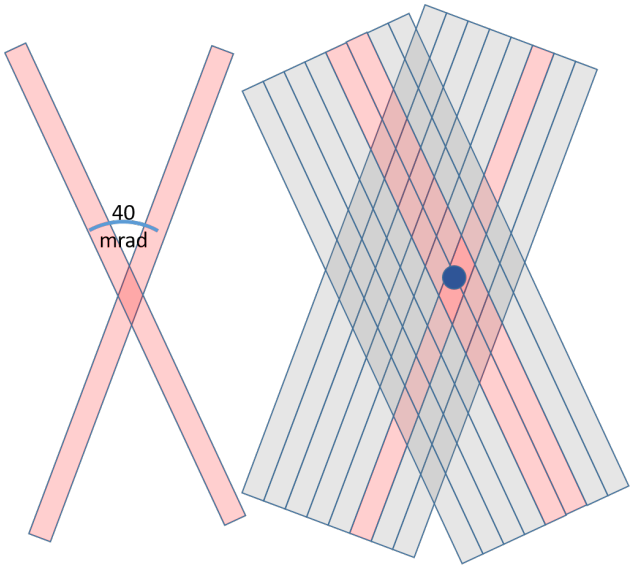

Figure 5.3: Schematic diagram of hit clusters on opposite sides of a given SCT module to produce a space-point.

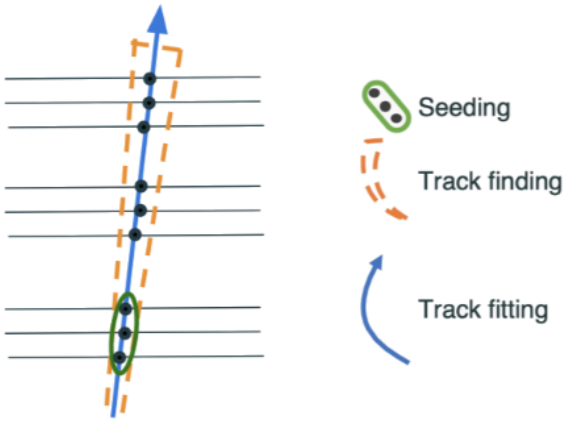

Figure 5.4: Schematic diagram of space points being used for track-finding and track-fitting using a Kalman filter.

5.1.4 Track-finding

After the space-points have been reconstructed in each tracking station, they can be combined to form *track seeds*, which are the linear track segments inside each tracking station. This

is illustrated in Figure 5.4, which diagrammatically illustrates both track seeding and track fitting, which will be discussed in the following subsection. As there is no magnetic field in this region, both the truth and reconstructed track segments will be linear.

The track segments are constructed by combinatorially selecting pairs hits in the first and third tracking planes in each tracking station, then creating a triplet by adding one additional space point in the remaining tracking plane.

5.1.5 Track-fitting with the combinatorial Kalman filter

Once the track seeds have been reconstructed, the final task in tracker reconstruction is to combine them to form track candidates. This is done by accounting for the helical track segment in the dipole volume between adjacent tracking stations which implies there will be a relative misorientation of track seeds in different tracking stations.

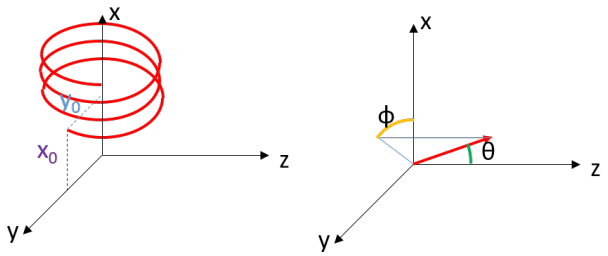

Figure 5.5: Schematic of helix track parameters.

ACTS tracks in the FASER spectrometer are described by five track parameters, namely $(x_0, y_0, \theta, \phi, q/p)$, where (x_0, y_0) are the transverse coordinates of the track in the first layer, θ is the angle of the start of the track with the z-axis, ϕ is the angle in the (x, y)-plane of the start of the track, and q/p is the electron charge divided by the momentum.

This track-fitting is done using a *combinatorial Kalman filter*. The track parameters to be fitted form a vector $\vec{x} = (x_0, y_0, \theta, \phi, q/p)$. As the particle traverses the FASER spectrometer, this vector is updated. Let $\vec{x}_k$ denote the track parameters at tracker plane k.

What are actually measured by FASER are not the track parameters but measurements $\vec{m}_k$. This is related to the track parameters by

$$\vec{m}_k = \vec{h}_k(\vec{x}_k) + \vec{\epsilon}_k, \tag{5.1}$$

where $\vec{h}_k$ is the measurement function and $\vec{\epsilon}_k$ the errors in measurement, with expectation values of $\langle \vec{h}_k \rangle = \vec{m}_k$ and $\langle \vec{\epsilon}_k \rangle = \vec{0}$.

The state vector $\vec{x}_k$ evolves according to

$$\vec{x}_k = \vec{f}_k(\vec{x}_{k-1}) + \vec{w}_k, \tag{5.2}$$

where $\vec{w}_k$ represents the error inherent in the fluctuating track parameters themselves and also has expectation $\langle \vec{w}_k \rangle = \vec{0}$.

In the immediate neighbourhood of $\vec{x}_{k-1}$ the function $\vec{f}(\vec{x}_{k-1})$ can be linearized as $\mathbf{F}_k \vec{x}_{k-1}$ where $\mathbf{F}_k$ is now a square matrix; likewise $\vec{h}_k(\vec{x}_k) \approx \mathbf{H}_k \vec{x}_k$. This gives a linear dynamic system. The system equation is

$$\vec{x}_k = \mathbf{F}_k \vec{x}_{k-1} + \vec{w}_k, \tag{5.3}$$

and the measurement equation is

$$\vec{m}_k = \mathbf{H}_k \vec{x}_k + \vec{\epsilon}_k. \tag{5.4}$$

Denote by $\mathbf{Q}_k(\mathbf{V}_k)$ the covariances of $\vec{w}_k(\vec{\epsilon}_k)$.

The Kalman filter then proceeds as follows:

- The state vector is extrapolated from $k-1$ to k using the estimate $\vec{x}_k^{k-1} = \mathbf{F}_k \vec{x}_{k-1}$.
- The measurements are used to update the state vector using $\vec{r}_k^{k-1} = \vec{m}_k - \mathbf{H}_k \vec{x}_k^{k-1}$.

In effect, each step of the Kalman filter predicts the state vector at the next step, then updates it using the measurements at that step. The Kalman filter uses the measured hits in each layer to determine the track parameters, then updates those parameters based on the measured hits in the next layer. This iterative procedure determines the charged tracks in FASER using all available space points.

5.2 Estimation of tracker performance

It is important to perform a sanity check to verify that the tracker performance is consistent with expectations. This requires a back-of-the-envelope calculation which relies on the cyclotron motion exhibited by charged particles passing through a magnetic field.

The radius of the trajectory of the particle (including relativistic effects) is given by $R = p_T/(qB)$ where q is the charge of the particle and B the magnetic field. If L denotes the distance between the first and last sensitive detector element, then the angleθ subtended by the trajectory at the center satisfies

$$2\sin\frac{\theta}{2} = \frac{L}{R}. \tag{5.5}$$

For high-energy particles the arc is approximately straight and the angle small, so the approximation $\theta \approx L/R$ is appropriate.

An important parameter is the *sagitta,* or transverse distance between the midpoint of the arc and the chord connecting its endpoints. The sagitta s is given by

$$s = R\left(1 - \cos\frac{\theta}{2}\right) \approx \frac{R\theta^2}{8}. \tag{5.6}$$

Since $\theta \approx L/R$ and $R \approx p_T/(qB)$, it follows that $\theta \approx qBL/p_T$ and

$$s \approx \frac{p_T}{qB}\frac{q^2B^2L^2}{8p_T^2} = \frac{qBL^2}{8p_T}. \tag{5.7}$$

The fractional momentum resolution is then given by the fractional sagitta resolution:

$$\frac{\sigma(p_T)}{p_T} \approx \frac{8p_T}{qBL^2}\sigma_s. \tag{5.8}$$

Note that the fractional momentum resolution grows linearly with momentum due to the intrinsic resolution. For the FASER detector with four tracking stations this is

$$\frac{\sigma(p_T)}{p_T} \approx (5.272 \times 10^{-4})\left(\frac{p_T}{\mathrm{GeV/c}}\right). \tag{5.9}$$

At $p_T = 500\,\mathrm{GeV}/c$, this is $\sigma(p_T)/p_T \approx 0.26$, in rough agreement with the 0.18 found above.

A similar calculation can be carried out for the FASER spectrometer with three tracking stations only (i.e., no IFT). The derivation is exactly the same, but the total detector length L is reduced by a factor of 4200/2500 to account for the shortening of the sensitive tracking volume. This gives

$$\frac{\sigma(p_T)}{p_T} \approx (1.487 \times 10^{-3}) \left(\frac{p_T}{\mathrm{GeV/c}} \right). \tag{5.10}$$

This is a threefold worsening of the resolution compared with Equation 5.9.

Chapter 6

ELECTRON CLUSTERING AND ENERGY RECONSTRUCTION

6.1 Muon Subtraction

Muons passing through the FASERν pilot detector will travel in very straight tracks due to their high mass and energy. As a result, in the simulation and pilot dataset, it is necessary to identify basetracks that belong to straight tracks, as straight tracks are a necessary (though not sufficient) condition to identify muons in the pilot detector.

The emulsion tracking algorithm proceeds as follows:

1. Basetracks from each emulsion film are selected according to the cuts $Tx - cTx < 10\,\text{mrad}$ and $Ty - cTy < 10\,\text{mrad}$, where cTx and cTy are the average values of Tx and Ty evaluated for all basetracks in the emulsion film.

2. The quality-of-fit parameter χ^2 is determined for every pair of basetracks.

3. If $\chi^2 > 3.0$, the link between basetracks is rejected.

4. All basetracks with valid links are connected. At this stage basetracks may have links to multiple other basetracks that are in the same layer.

5. Links between basetracks (e.g., $A - C$) that are also connected by an intermediate series of links (e.g., $A - B - C$) are deleted.

6. The best-fit link is determined by constructing two triangles whose vertices are the positions of three basetracks and the differences between their relative displacements

and angles. The links which produce triangles with the smallest area are selected, and promoted to straight tracks.

7. The deletion of links in Steps 5-6 will have caused some tracks to be separated. These track segments are now reconnected by calculating χ^2 for basetracks in separate track segments and matching those with the lowest value of χ^2.

Steps 5-6 are illustrated by example in Figure 6.1. As shown, there are links $A-B$, $A-C$, $A-D$, $B-C$, $B-D$, and $C-D$. Because nodes A and C are linked through the sequence of links $A-B-C$, the link $A-C$ is deleted. Likewise, the nodes B and D are linked through the sequence of links $B-C-D$, so the link $B-D$ is deleted. Lastly, the nodes A and D are linked through the sequence of links $A-B-C-D$, so the link $A-D$ is deleted. Figure 6.2 shows the two triangles whose areas are minimized in order to find the

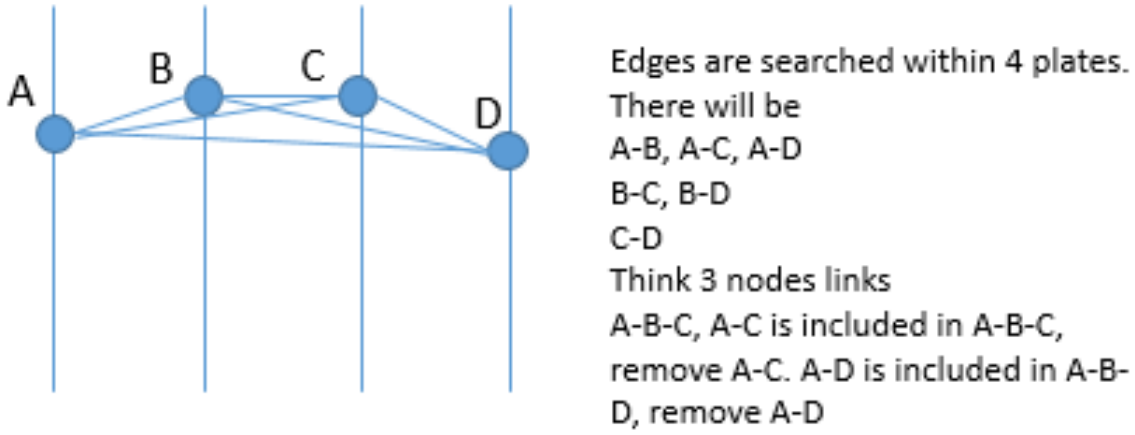

Figure 6.1: Illustration of four basetracks (nodes A-D) for step 5. The links between basetracks are deleted in the sequence shown at right.

most credible path between successive layers. These allow for the selection of one sequence of basetracks which are deemed most likely to belong to each straight track. Step 7 is illustrated in Figure 6.3. The left-most panel shows the fully linked set of basetracks, from

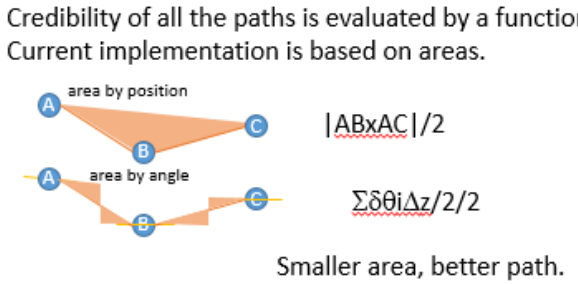

Figure 6.2: Visualization of the position (above) and angular (below) triangles whose areas are used to evaluate the credibility of a link.

which multiple connections are deleted and the credibility is evaluated using the areas of the triangular regions to select a most credible path. Finally, the tracks that had been separated by the deletion process are reconnected by evaluating χ^2 for basetracks in the two separate tracks. Now that basetracks have been linked to form straight tracks, three further

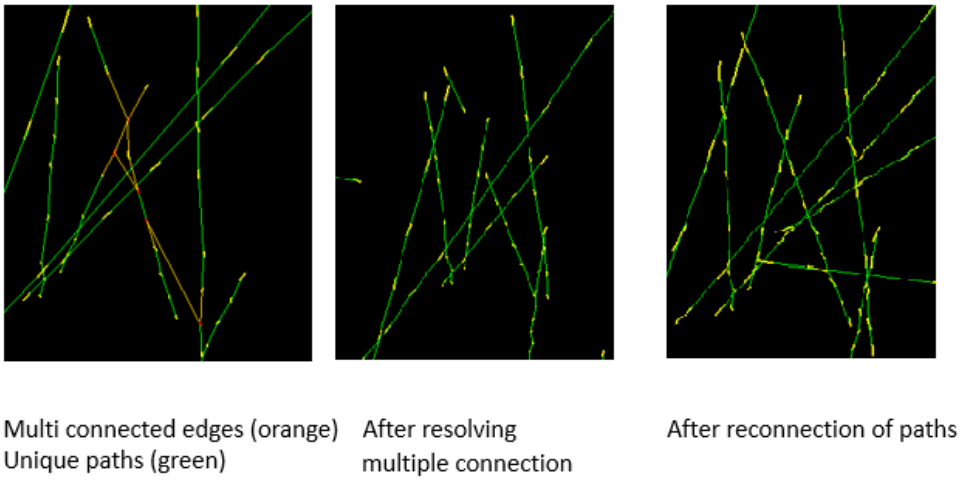

Figure 6.3: Example illustration of the track-linking procedure. After all suitable basetrack links are made (left), the procedure to remove multiple links is applied and the most credible path is selected (middle). The straight tracks are then reconnected (right).

conditions are imposed on the basetracks belonging to straight tracks:

- The first basetrack in the straight track must be within the first 4 layers of the detector.
- There must be at least 4 basetracks which constitute the straight track.
- The last basetrack must be at least 10 layers from the first basetrack in the same straight track.
- The basetracks of the straight track must be within 10 mrad of the mean of the distributions of Tx, Ty, i.e., within 10 mrad of $(-0.008, 0.028)$ rad.

The first three conditions follow from the property that muons are expected to pass through the entire pilot detector, while the last condition arises from the assumption of the origin of the muons at the ATLAS interaction point as well as the muon angular distributions.

After pre-processing and muon tracking, the next step is to subtract off basetracks corresponding to muons, leaving primarily e^+/e^- basetracks that can be used as input to the clustering algorithm. There exists a tracking algorithm, used by DsTau, for reconstructing straight tracks from basetracks. However, not all of the reconstructed straight tracks should be identified as muons. To determine more accurate muon identification criteria, it is noted that muons will tend to penetrate through many layers of the FASERν pilot detector and will tend to originate from the ATLAS interaction point. These observations inspire the following selection criteria:

- The reconstructed straight track contains at least four basetracks.
- The first basetrack in the reconstructed straight track occurs in the first six plates of the pilot data sample.
- The track angles in the (x, z)- and (y, z)-planes are within 0.01 rad of the angular center $(-0.008, 0.028)$ rad (offset from $(0, 0)$ due to manual alignment of the FASERν pilot detector).

Applying these three criteria leads to a track density of 1.9×10^5 penetrating tracks per square centimeter of pilot detector area, which agrees well with the simulation of 2.1×10^5 expected muons per square centimeter. This validates the penetrating track criteria.

After muon subtraction, the remaining basetracks are almost all e^+/e^- basetracks, either as knock-on electrons induced by the muons or as part of the EM shower initiated by it. This is the raw input to be used for electron clustering.

6.2 Electron clustering algorithm

In order to perform electron clustering efficiently, it is necessary to identify the seeds themselves. This is done in a two-step procedure, where the first step is run on all basetracks and the second step is run on the seeds identified in the first step.

- Perform electron clustering with the constraints $\alpha \leq 0.001$ rad, $\Delta r \leq 30\,\mu$m, and $\Delta\theta \leq 0.001$ rad and to a depth of five plates. If this cluster contains at least three basetracks in any plate, these likely correspond to the primary electron and an e^+/e^- pair, and should be identified as an electron seed.

- Perform electron clustering with the constraints $\alpha \leq 0.03$ rad, $\Delta r \leq 50\,\mu$m, and $\Delta\theta \leq 0.03$ rad and to a depth of 29 plates, starting with identified electron seeds.

This seeding algorithm was first used by Kobayashi in 2012 [35]. These criteria are motivated by the fact that true EM showers tend not to be conical, but are rather contained in a narrow cone before spreading out due to the higher energy (and therefore longitudinal momentum) of e^+/e^- basetracks close to the seed. The cuts are so stringent that the lower energy e^+/e^- basetracks further downstream do not pass the seeding criteria.

This electron clustering algorithm is illustrated schematically in Figure 6.4. From the seed basetrack in blue on the far left, a 30 mrad cone is constructed in orange and all basetracks within this cone are considered for clustering as shown in black in the top diagram. Basetracks in layer k are then extrapolated to layer $k+1$. For each of these selected basetracks,

the transverse position difference Δr and angular difference $\Delta\theta$ are calculated. Those selected basetracks which pass the clustering selection criteria $\Delta r < 50\,\mu$m and $\Delta\theta < 30\,$mrad are shown in black and added to the cluster; those that fail the criteria are shown in red and not added to the cluster. The procedure is then repeated, extrapolating all clustered basetracks in the most downstream layer and adding those basetracks in the cone which pass the clustering criteria, until there are no more basetracks to be added. Note that, in principle, all clusters should end because the produced electrons have progressively lower energies and therefore scatter more in the emulsion detector, so at some point the angular differences between basetracks belonging to one layer of the cluster and those belonging to the next layer will fail one or both of the clustering criteria.

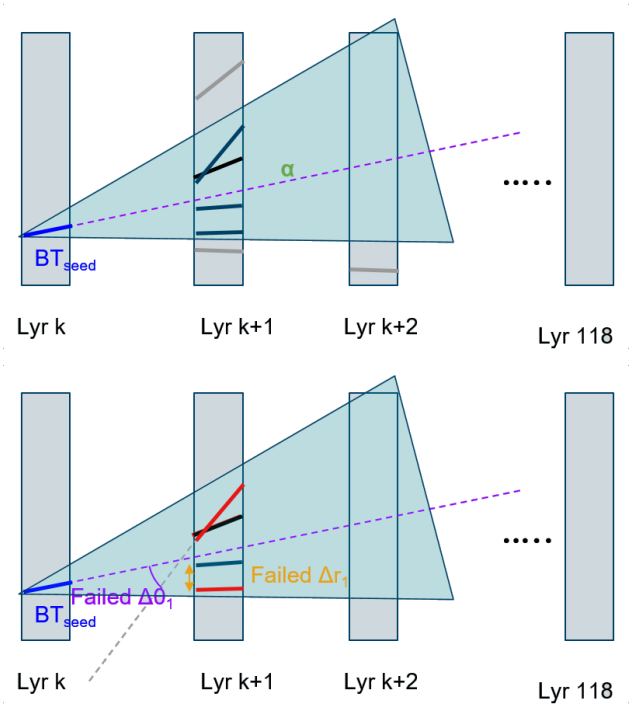

Figure 6.4: Diagram showing the electron clustering procedure, starting with the seed basetrack (blue), constructing the cone (green), selecting basetracks within the cone, extrapolating and applying cuts on Δr and $\Delta\theta$, and adding basetracks to the electron cluster (black). Basetracks which fail the matching criteria (red) are not clustered.

The reconstructed electron clusters produced by this procedure will be contaminated with basetracks from EM showers initiated by other electrons. To minimize the number of such basetracks incorrectly clustered by the clustering algorithm, the impact parameter (IP) of basetracks with respect to the first plate in the cluster is considered. Signal basetracks (namely, those originating from the same electron as produced the seed basetrack) tend to have smaller IPs, while background basetracks belonging to showers initiated by other electrons will generally have larger IPs. From a simulation study, the cut $IP < 250\,\mu$m reduces the number of background basetracks to a negligible amount while still preserving a large fraction of the signal EM shower.

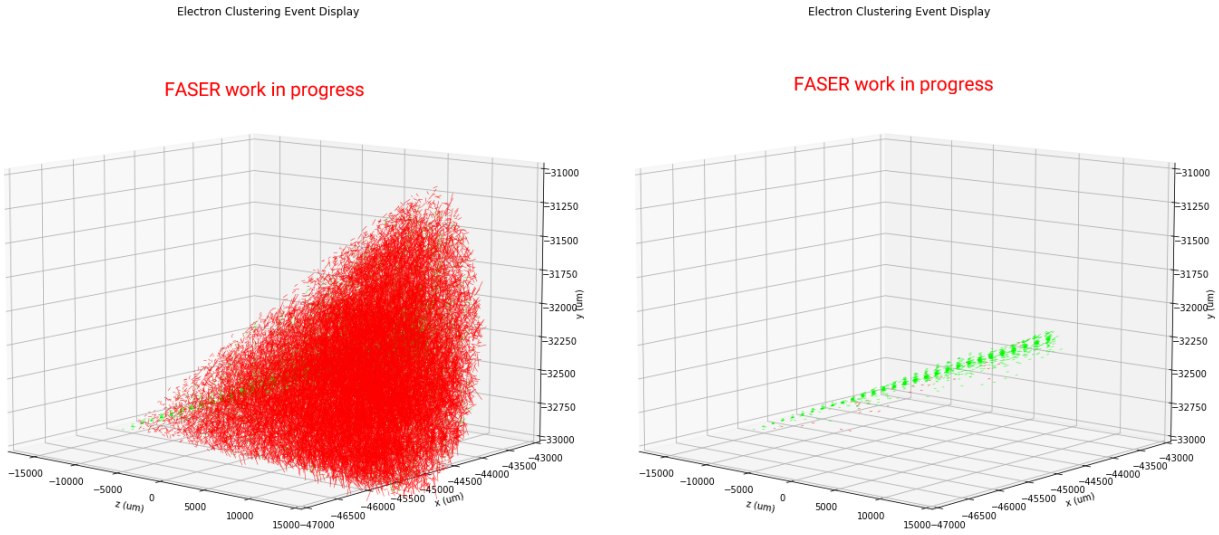

Figure 6.5: Example reconstruction of electron cluster with background contamination from other clusters (left), and clustered electron after purification (right).

However, when clustering the simulation data without the truth-cluster instance segmentation, it is immediately apparent from the left panel of Figure 6.5 that many basetracks belonging to other knock-on electron clusters are incorrectly associated to the same reconstructed cluster. Basetracks belonging to an incorrect truth cluster are shown in red, while

basetracks belonging to the correct truth cluster are shown in green. Figure 6.6 shows the

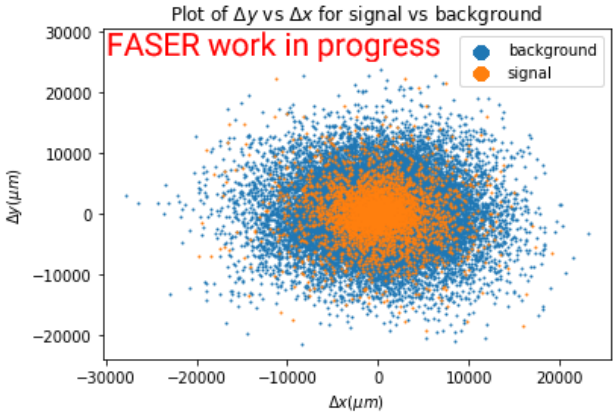

Figure 6.6: Plot of projections of clustered basetracks to emulsion film of seed basetrack.

linear extrapolation of all clustered basetracks back to the first emulsion film in the electron cluster. The clustered basetracks which belong to the correct truth cluster have extrapolated positions that are closer to the seed basetrack position than are the extrapolated positions of clustered basetracks belonging to incorrect truth clusters. After applying the cut $IP < 350\,\mu$m, the signal-background separation improves dramatically, as shown in the right-hand panel of Figure 6.5.

Figure 6.7 shows 29 reconstructed electron clusters. It is clear that the IP cut is effective in separating signal from background, and therefore it is expected that most clustered electron basetracks do indeed belong to the same truth cluster.

This gives the following procedure for clustering the knock-on electrons and purifying the resulting clusters:

1. Construct a 100 mrad cone about the seed basetrack. Only tracks within this cone will be considered for clustering.

2. Extrapolate the seed basetrack to the subsequent emulsion film.

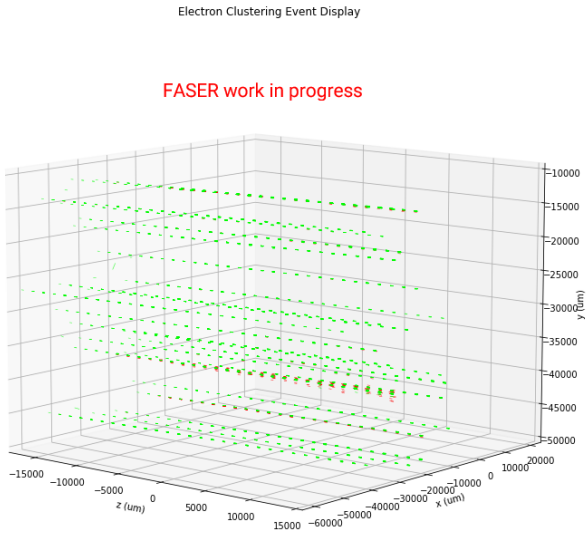

Figure 6.7: Event display of multiple reconstructed electron clusters purified with IP cut.

3. Add to the electron cluster any tracks in the film which are within 50 μm and 10 mrad of the extrapolated seed track.

4. Iterate the process, extrapolating all clustered basetracks in emulsion film i to emulsion film $i + 1$, and adding those which meet the above matching criteria to the cluster.

5. When there are no more basetracks matching any extrapolated tracks from the previous emulsion film within the matching criteria, end the clustering procedure.

6. Remove from the cluster any basetracks with $IP > 350\,\mu$m.

6.3 Electron energy regression

The final step is to count the number of basetracks in the purified cluster and use this to reconstruct the electron energy. There are three variables available for electron energy regression:

- The total number of basetracks;
- The maximum track multiplicity in one layer;
- The depth of the shower maximum.

The fundamental principle is that the number of truth basetracks produced in an electron EM shower is approximately proportional to the energy of the primary electron. This trend persists with clustered basetracks instead of truth basetracks, although the trend is weaker than the truth-level results present in Appendix C. Figure 6.8 shows two variables used

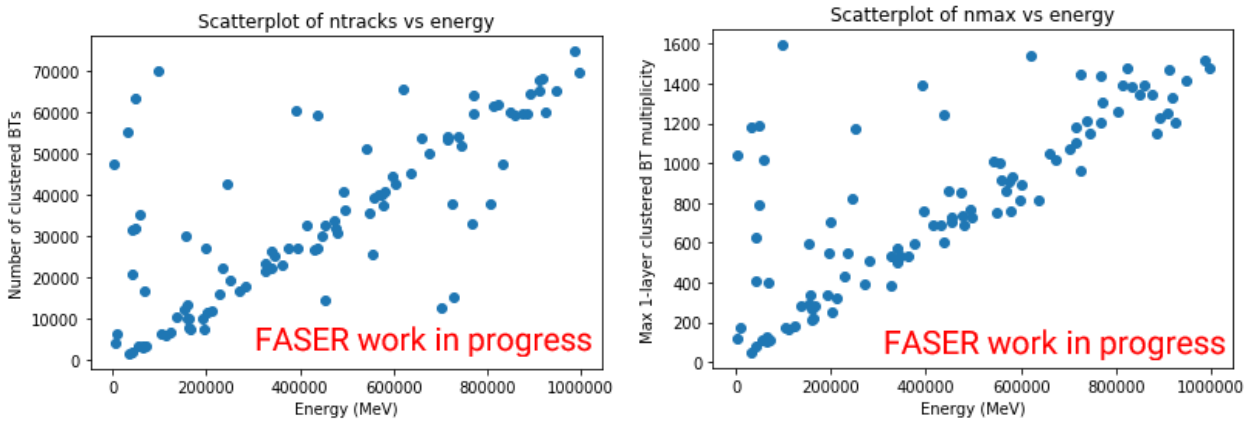

Figure 6.8: Plots of total number of clustered basetracks N_{total} (left) and maximum clustered basetrack multiplicity in a single layer N_{max} (right) as functions of energy.

for electron energy regression: The total number of clustered basetracks N_{total}, and the maximum multiplicity of clustered basetracks within a single layer N_{max}. Observe the strong linear trend of both with energy. At very low energies, the number of clustered basetracks and the maximum basetrack multiplicity both have a large deviation. This is indicative of the fact that at these energies, the purification algorithm is not particularly effective in separating signal from background and, therefore, many clusters at low energies have high basetrack counts due to low purity. Figure 6.9 shows the depth of the shower maximum iz_{max} versus the incident electron energy. There is a significantly weaker trend present in

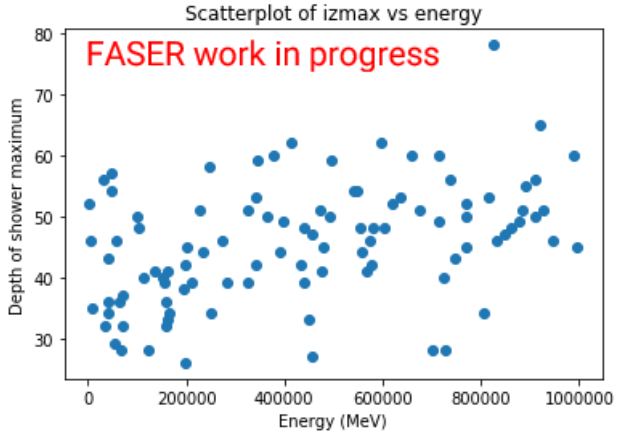

Figure 6.9: Plot of the depth (emulsion plate index) of the shower maximum $iz_{\max}$ as a function of energy.

this dataset, meaning that the expected predictive power of this variable is much lower than that of N_{total} and $N_{\max}$.

Chapter 7

PILOT RUN ANALYSIS

For the pilot run in 2018, $12.2\,\mathrm{fb}^{-1}$ of data were collected with the two FASERnu pilot detector modules. This pilot run has recorded the first neutrino interaction candidates at the LHC, as reported in [33]. There is a background of muons incident on the pilot detector at a density of approximately 2.1×10^5 muons per square centimeter, with the charge distribution being roughly even between positive and negative muons.

For each basetrack recorded in the FASERnu pilot detector, there are two angles of interest, namely $\theta_x = \arctan(p_x/p_z)$ and $\theta_y = \arctan(p_y/p_z)$. What is recorded by the emulsion detector is not θ_x and θ_y, but rather $Tx = \tan\theta_x$ and $Ty = \tan\theta_y$. These are the slopes the basetrack makes with the z-axis in the (z, x)- and (z, y)-planes, respectively. It is clear that the angular distributions of basetracks should peak at $Tx = Ty = 0$, since the basetracks will be recorded by particles originating from primary particles produced at the ATLAS interaction point 470 meters away, and the FASERnu pilot detector is only $12.5\,\mathrm{cm} \times 10\,\mathrm{cm}$ in area.

The study of interest with the FASERnu pilot run is that of electron energy reconstruction in the pilot detector. The reasons for this are twofold:

- The energy spectrum of positive and negative muons incident on the FASERnu pilot detector from the ATLAS interaction point has been obtained via dedicated simulations. However, the actual data from the FASERnu pilot run allows for the validation of the simulation spectrum. The high-energy muons incident on the pilot detector will occasionally interact with atoms to produce a secondary electron (so-called knock-on electrons or δ-rays). By simulating the expected muon flux incident on the FASERnu pilot detector, it is possible to identify and reconstruct the EM showers of these knock-

on electrons and therefore their energies. By applying the same procedure to the pilot run data and pilot run MC simulation, the number of identified knock-on electrons can be compared between data and simulation, and the energy spectra compared. If the knock-on electron energy spectra are consistent, the muon energy spectrum will be validated.

- In order to reconstruct the energy of an electron neutrino CC event, the energy of the outgoing electron must be reconstructed. It is known that higher energy electrons will leave larger clusters (that is, clusters containing more e^+/e^- basetracks), and in particular the number of truth e^+/e^- basetracks in an electron EM shower grows approximately linearly with energy. A well-designed electron clustering algorithm should produce a similar linear relationship between clustered basetracks and truth electron energy. Using this, we can reconstruct the electron energy and thus the neutrino energy by counting the number of basetracks associated to the cluster, and therefore measure the electron neutrino CC cross section by counting the number of electron neutrino interactions in FASERnu in each energy bin.

The data recorded by the FASERν pilot detector is stored in the FEDRA format, containing the following information:

- Position (x, y, z), angular $(\tan\theta_x, \tan\theta_y)$, and quality-of-fit for microtracks used for emulsion reconstruction.

- Position, angular, and quality-of-fit for basetracks reconstructed from the above microtracks.

These data are stored in 30 subvolumes for the pilot detector, where each subvolume is a square cylinder with cross-section 22 mm $\times$ 22 mm and overlaps with neighbouring subvolumes by 2 mm. For the electron clustering study, the last 29 (highest-quality) plates are used.

The first step is to pre-process the raw emulsion data. These data are stored by individual emulsion plates and record local positions relative to the plate; by reading these files in with

FEDRA, the basetracks are aligned in both the transverse and the longitudinal directions, as necessary to prepare the data sample for clustering.

Over the 29 plates considered for this study, the electron cluster may spread out to a diameter of 7 mm. Since there is only an overlap of 2 mm between adjacent subvolumes, it is clear that up to four subvolumes in a 2×2 block must be considered in order to reconstruct the entire electron cluster. Moreover, the 2×2 subvolume blocks must overlap so that an electron cluster which escapes one 2×2 block is contained entirely within another.

7.1 Comparison of raw basetrack parameters between MC and data

Figure 7.1 shows a comparison of the distributions of the basetrack position variables x, y, z as well as the angular variables Tx and Ty between the FASERnu pilot data and the MC simulation. The normalization of the MC simulation dataset is based on the number of expected muons ($\approx 2 \times 10^7$ over the entire pilot detector volume), which is based on the result from the FASERν pilot run data. The position distributions are approximately consistent, providing a first validation of the MC simulation.

For distributions of the variables Tx and Ty, note that there is an offset in both variables, so that their peaks occur at approximately $\langle Tx \rangle = -0.008$ and $\langle Ty \rangle = 0.028$, respectively. This difference from the expected mean values indicates misalignment of the FASERnu pilot detector in the TI18 maintenance tunnel with the beam axis. This observed misalignment is not surprising because the alignment of the FASERnu pilot detector was performed manually without precision metrology equipment. The GEANT4 simulation of both μ^+ and μ^- with the expected energy and angular distributions was then prepared using these angular offsets.

7.2 Muon identification

The muons passing through the pilot detector are distinguished from the knock-on electrons and their corresponding EM showers by the straightness of their tracks; whereas electrons moving through the pilot detector will scatter frequently and lose energy rapidly, muons will pass almost unaffected through the pilot detector on account of their mass being approxi-

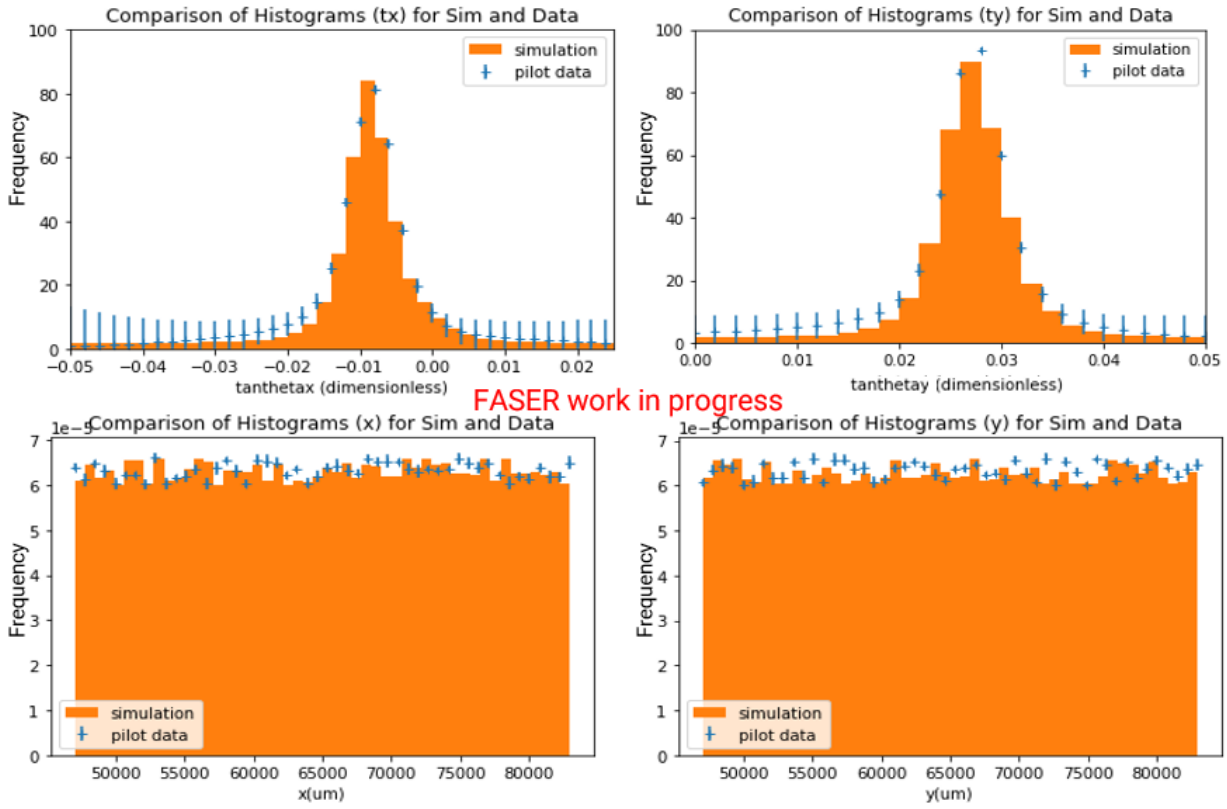

Figure 7.1: (top-left) Histogram of x for emulsion hits in truth MC simulation superposed with x for basetracks in one 2×2 subvolume block of the pilot run data samples. (rop-right) Histogram of y for emulsion hits in truth MC simulation superposed with y for basetracks in one 2×2 subvolume block of the pilot run data samples. (bottom-left) Histogram of $\tan\theta_x$ for emulsion hits in truth MC simulation superposed with $\tan\theta_x$ for basetracks in pilot run data samples. (bottom-right) Histogram of $\tan\theta_y$ for emulsion hits in truth MC simulation superposed with $\tan\theta_y$ for basetracks in pilot run data samples.

mately 207 times greater. Therefore basetracks which align sufficiently well with a straight track, and which penetrate through multiple layers of the pilot detector, are identified as likely muon candidates.

The algorithm for identifying basetracks belonging to straight linked tracks evaluates the quality of links between basetracks using the parameter χ^2, defined as

$$\chi^2 = \frac{\Delta\theta_L^2}{\sigma_L^2} + \frac{\Delta\theta_T^2}{\sigma_T^2} + \frac{\Delta x^2}{\sigma_x^2} + \frac{\Delta y^2}{\sigma_y^2},$$

where $\Delta\theta_L$ and $\Delta\theta_T$ are the longitudinal and transverse angular differences between the two basetracks. Δx and Δy are calculated for the two basetracks by extrapolating both linearly to the midpoint of the z position between them and calculating the differences in the x- and y-coordinates in this shared plane. The transverse angular resolution of a basetrack is determined from the position resolution of the microtracks making up the basetrack, $\sigma_{MT,x} = 0.3\,\mu$m, and the thickness of the plastic base between the emulsion layers recording the microtracks, $T_{\text{base}} = 100\,\mu$m, by

$$\sigma_T = \sqrt{(5\,\text{mrad})^2 + \left(\frac{\sigma_{MT,x}^2}{T_{\text{base}}^2}\right)} = \sqrt{(5mrad)^2 + (3mrad)^2} \approx 5.8\,\text{mrad}$$

and the longitudinal angular resolution is given in terms of the longitudinal microtrack position resolution $\sigma_{MT,z} = 5\,\mu$m by

$$\sigma_L(Tx, Ty) = \sqrt{\sigma_T^2 + (Tx^2 + Ty^2)\frac{\sigma_{MT,z}^2}{T^2}} = \sqrt{(5.8\,\text{mrad})^2 + (Tx^2 + Ty^2)(0.05\,\text{mrad})^2}.$$

The longitudinal angular resolution manifestly depends on the slopes Tx and Ty, as expected; basetracks with smaller slopes are expected to match more precisely than basetracks with larger slopes due to the greater distance separating the basetracks.

The resolutions σ_x, σ_y are determined by

$$\sigma_x = \sqrt{\sigma_{x0}^2 + (800\,\mu\text{m})\sigma_L^2}, \qquad \sigma_y = \sqrt{\sigma_{y0}^2 + (800\,\mu\text{m})\sigma_L^2},$$

where $\sigma_{x0}^2 = \sigma_{y0}^2 = (2\,\mu\text{m})^2$ and the effects of the angular resolution have been taken into account.

The emulsion tracking algorithm then proceeds as outlined in Section 6.1. Note that the FASERnu pilot data is divided into 30 subvolumes in a 5×6 grid pattern. Each subvolume covers a 26 mm $\times$ 26 mm area and has a 6 mm overlap with each of its neighbouring regions. Because of the division into subvolumes, it is possible that straight tracks may leave a subvolume before exiting another. However, by selection, the basetracks are at most 0.038 rad from normal incidence and propagate through 29 layers (≈ 35 mm) of the pilot detector, meaning

they can travel a transverse distance of at most 1.4 mm. Due to this steep incidence of muon-induced straight track, as well as the 6 mm overlap with the neighbouring subvolume, those tracks that exit one subvolume will be entirely contained in the neighbouring subvolume.

After applying these conditions to the basetracks which constitute straight tracks, a total of 3.0×10^7 penetrating tracks are found over the pilot detector volume. This is consistent with the 2.7×10^7 muons expected to be incident on the pilot detector, a second validation of the simulation.

7.3 Electron clustering sample preparation

The next step is to subtract from all basetracks (in the FASERnu pilot run data and the simulation) those basetracks belonging to penetrating straight tracks. The surviving basetracks are likely to be e^+/e^- basetracks, produced either as knock-on electrons from the incident muons or as part of an EM shower initiated by these electrons.

Figures 7.2-7.5 show the evolution of basetracks during the muon subtraction process, in a $(200\,\mu\text{m})^2$ region centered at $(50, 50)$ mm with randomly coloured basetracks for clear separation of the basetracks involved. Starting from all basetracks, as shown in Figure 7.2, the muon tracking algorithm is executed to produce straight tracks as shown in Figure 7.3. Those basetracks which are part of straight tracks that also satisfy the penetrating-track criteria are shown in Figure 7.4, and all remaining basetracks that survive the subtraction are shown in Figure 7.5. These are almost all e^+/e^- tracks, and so will be used for the subsequent electron clustering.

The following electron clustering algorithm was developed for these high-energy electrons:

1. Start with a basetrack selected in layer i. (For $i = 90$ this is the seed basetrack.)

2. Construct a cone of apex angle $\alpha = 100$ mrad, with apex at the selected basetrack and axis of symmetry aligned with the basetrack.

3. Select all basetracks in the pilot detector which are within this 100 mrad cone. (This

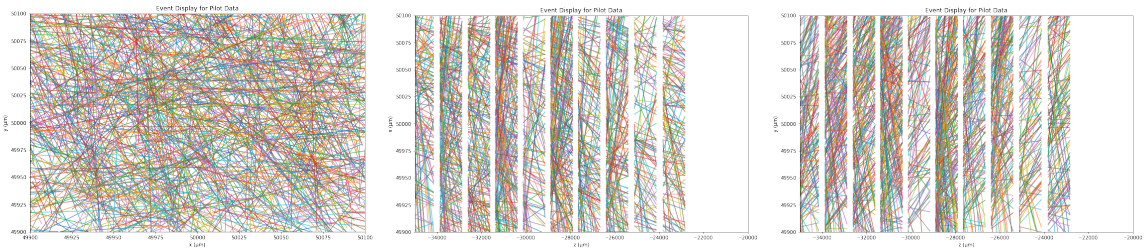

Figure 7.2: Two-dimensional event display showing xy- (left), xz- (middle), and yz-projections (right) of all basetracks within $(200\,\mu\text{m})^2$ region centered on $(x, y) = (50, 50)\,\text{mm}$. The lengths of the basetracks has been extrapolated to 1 mm to show the relationship between basetracks in successive layers.

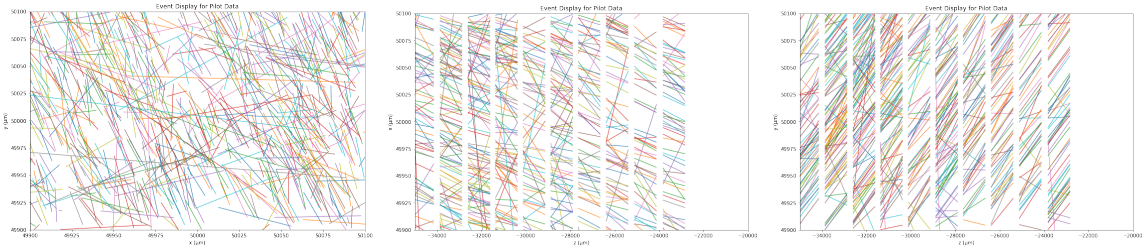

Figure 7.3: Two-dimensional event display showing xy- (left), xz- (middle), and yz-projections (right) of all basetracks that are components of reconstructed straight tracks within $(200\,\mu\text{m})^2$ region centered on $(x, y) = (50, 50)\,\text{mm}$.

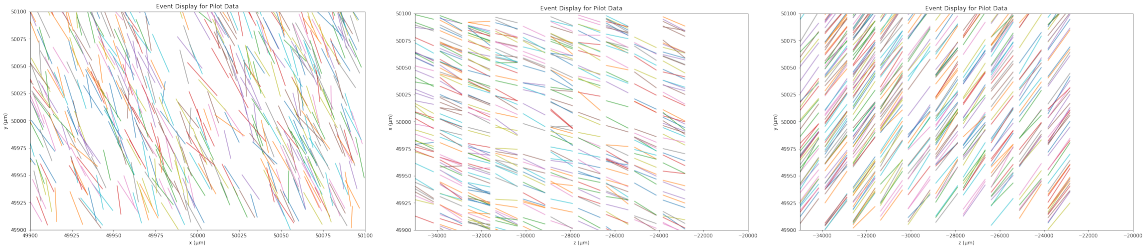

Figure 7.4: Two-dimensional event display showing xy- (left), xz- (middle), and yz-projections (right) of all basetracks that are components of reconstructed straight tracks within $(200\,\mu\text{m})^2$ region centered on $(x, y) = (50, 50)\,\text{mm}$ and that have passed the further cuts to be classified as penetrating tracks.

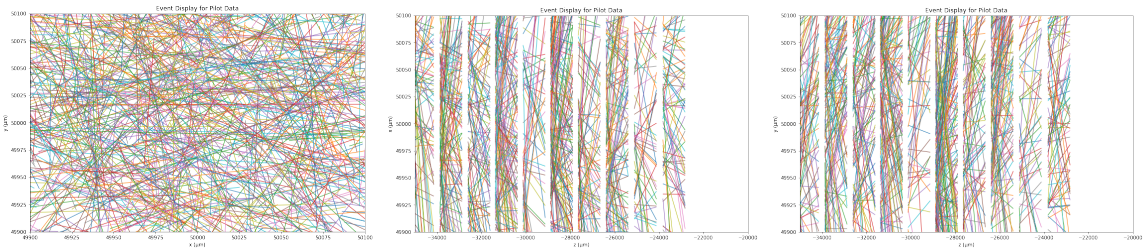

Figure 7.5: Two-dimensional event display showing xy- (left), xz- (middle), and yz-projections (right) of all basetracks within $(200\,\mu\text{m})^2$ region centered on $(x, y) = (50, 50)\,\text{mm}$ that are not part of penetrating straight tracks.

reduces the computational complexity of the problem and accelerates the algorithm.)

4. From the basetracks selected in Step 3, select those basetracks that are also in layer i+1.

5. Linearly extrapolate all clustered basetracks in layer i to layer i+1.

6. For each pair of basetracks in layer i and layer i+1, calculate the transverse distance Δr and angular difference $\Delta\theta$ between the extrapolated basetrack from layer i and the basetrack in layer i+1. If $\Delta r < 50\,\mu\text{m}$ and $\Delta\theta < 10\,\text{mrad}$, add the basetrack in layer i+1 to the cluster.

7. Repeat the process until no more basetracks are found.

8. When all clusters with seed basetracks in layer 90 have been clustered, subtract all clustered basetracks and repeat with seed basetracks in layer 91.

Note that basetracks are selected in a 100 mrad cone which may extend over 29 layers. In the final layer, the cross-section of the cone is described by a circle whose radius is approximately 5.5 mm, meaning that one shower may spread over a maximum of four subvolumes. Therefore it is necessary to combine the surviving e^+/e^- basetracks from four neighbouring subvolumes for electron clustering.

Figure shows the result of running the electron clustering algorithm on a 20 mm × 20 mm subvolume of the pilot detector centered on $(x, y) = (50, 50)$ mm. It is abundantly clear that there is overlap between electron clusters, with some basetracks included in multiple clusters. To reduce basetracks being clustered back to the wrong electron, it is important to note that

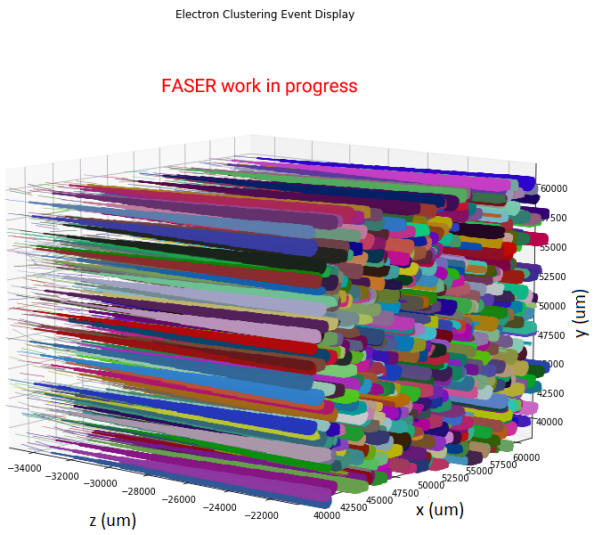

Figure 7.6: Event display of all clustered electrons within $(20\,\text{mm})^2$ area centered on $(x, y) = (50, 50)$ mm.

every basetrack recorded by the FASERν pilot detector has a unique basetrack identifier. This identifier will be the same even if the same basetrack is stored in files corresponding to different subvolumes. This allows for the unambiguous joining of different subvolumes. Because EM showers may extend over up to four different subvolumes in a 2×2 square configuration, $4 \times 5 = 20$ such combined subvolumes must be considered for electron EM shower clustering.

After running the electron clustering algorithm on the full pilot data sample, the algorithm clusters 112632 of 156613 truth electrons with energies above 100 GeV for an electron clustering efficiency of 71.9%.

7.4 Cluster purification in the pilot data

It is inevitable that basetracks registered by e^+/e^- belonging to different electron EM showers will be clustered together. The clusters need to be purified so that clustered basetracks belonging to different EM showers are removed. This can be accomplished using the impact parameter, calculated by extrapolating the basetrack in question to the layer containing the seed basetrack. Figure 7.7 shows the distribution of impact parameters for basetracks in the same reconstructed cluster which belong to different truth clusters. Clustered basetracks which belong to the same truth cluster have impact parameters that have a large peak in a region close to 0, whereas clustered basetracks which belong to a different truth cluster have impact parameters with a small peak that is offset from 0.

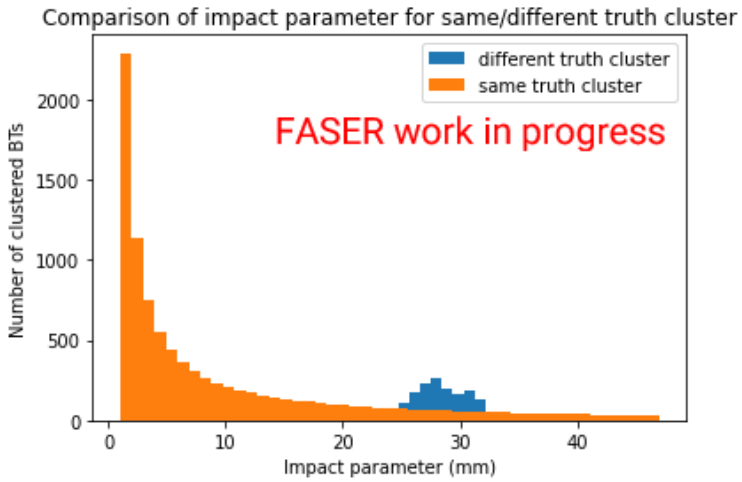

Figure 7.7: Distribution of impact parameters for clustered basetracks belonging to the same truth cluster as the seed basetrack (orange) and clustered basetracks belonging to a different truth cluster from the seed basetrack (blue).

After purifying the e^+/e^- clusters, the number of clustered, purified e^+/e^- tracks can now be counted as a function of the primary electron energy using the simulation. Figure 7.8 shows a plot of the number of basetracks associated to the electron cluster as a function of the truth electron energy, and there is clear indication of a linear trend.

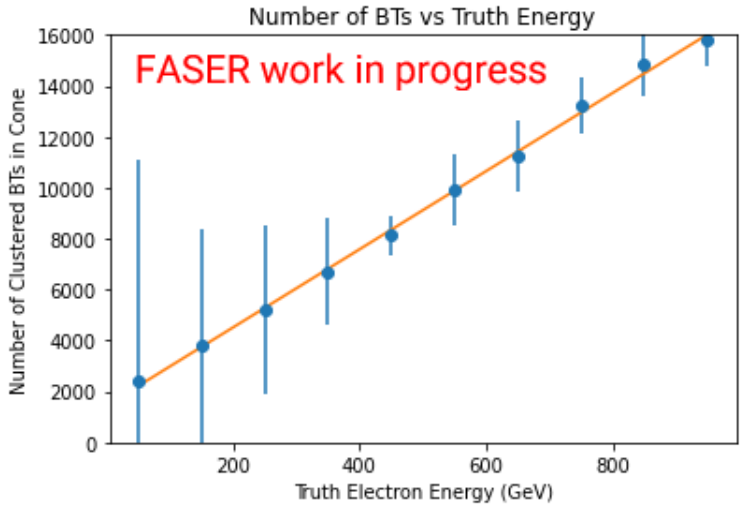

Figure 7.8: Scatter plot of number of clustered e^+/e^- basetracks as a function of the truth electron energy (blue points) along with a fit (orange line). Note that the number of basetracks grows approximately linearly with energy.

7.5 Electron energy regression with the pilot data

The electron energy resolution is plotted as a function of truth electron energy for the simulated pilot run sample in Figure 7.9. The resolution for truth electron energies below $E \approx 300\,\text{GeV}$ are very poor and so the reconstruction is unreliable in this region; however, at higher energies the electron energy regression has good resolution and so reconstruction of electron energies above 300 GeV are expected to be reliable.

Figure 7.10 shows a comparison of the number of e^+/e^- basetracks per layer for the pilot run simulation. We note that, while the number of clustered basetracks per layer is lower than the number of truth basetracks per layer, the distribution is qualitatively the same with the depth of shower maximum in the same place.

The distribution of muon energy losses and knock-on electron energies is shown in Figure 7.11, where muons producing knock-on electrons with energies $E > 10\,\text{GeV}$ are shown in blue and those producing knock-on electrons with energies $E > 100\,\text{GeV}$ are shown in orange. Note that the energies of produced knock-on electrons are similar to, but lower than,

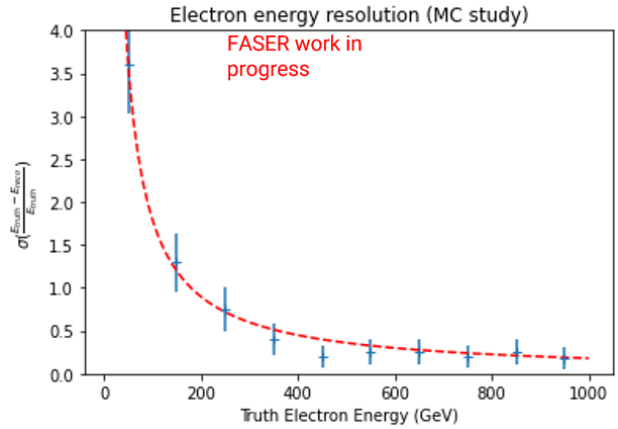

Figure 7.9: Electron energy resolution for one-dimensional energy regression using the total number of clustered, purified basetracks N_{total}. Note that below $E_{\text{truth}} \approx 300\,\text{GeV}$ the resolution is very poor and the results unreliable.

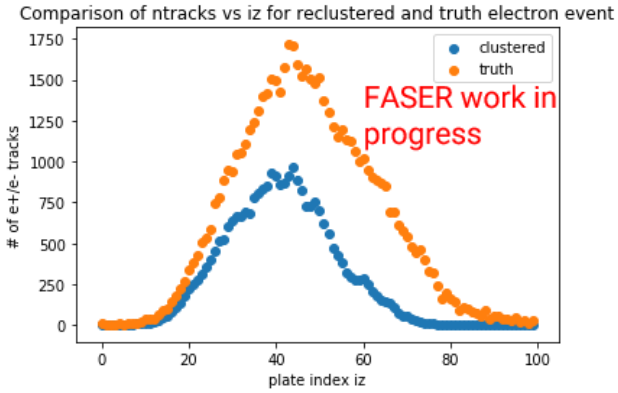

Figure 7.10: Comparison of number of truth basetracks per layer (orange) and number of clustered basetracks per layer (blue).

the energy losses of the parent muons, as expected.

Figure 7.12 shows the distribution of opening angles between produced knock-on electrons and the parent muons. Note that for events producing high-energy knock-on electrons with $E > 100\,\text{GeV}$, the opening angle is very narrowly distributed within approximately 1 mrad. The implication of this is that, because electron showers can potentially reach radii of $O(\text{cm})$, the parent muon will pass directly through the middle of the knock-on electron EM shower.

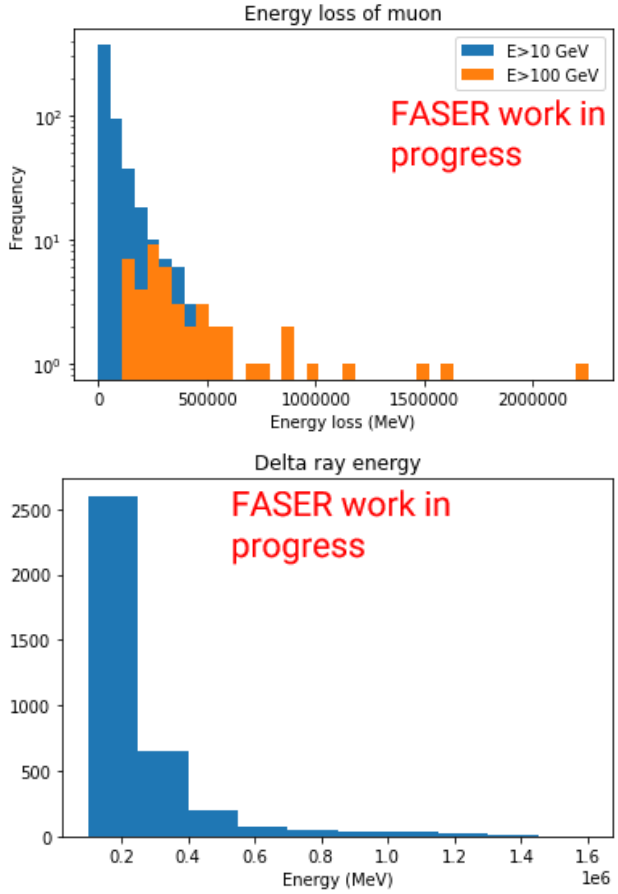

Figure 7.11: Distribution of muon energy losses (left) and distribution of energies of produced knock-on electrons (right) in the pilot run simulation. Note that the electron energy distribution is shifted to the left as compared with the muon energy loss distribution.

Now that an electron energy regression model has been developed, it is applied equally to the GEANT4 simulation and to the identically-formatted FASERnu pilot run data. The electron energy regression was performed on simulation and data samples taken from subvolumes in a 2×2 overlapping arrangement. The resulting comparison plot shows good agreement between simulation and data, as can be seen in Figure 7.13.

It is important to validate the GEANT4 simulation of the FASERν pilot detector. To this end, a study was done to compare the simulated data for the pilot detector with the

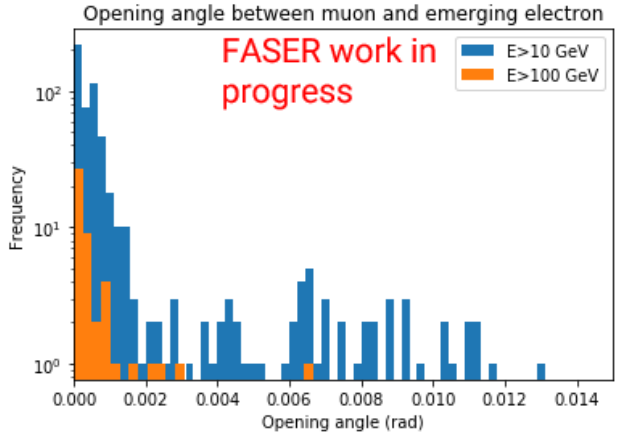

Figure 7.12: Histogram of opening angle between knock-on electron and parent muon for electrons with energies $E > 10\,\text{GeV}$ (blue) and $E > 100\,\text{GeV}$ (orange).

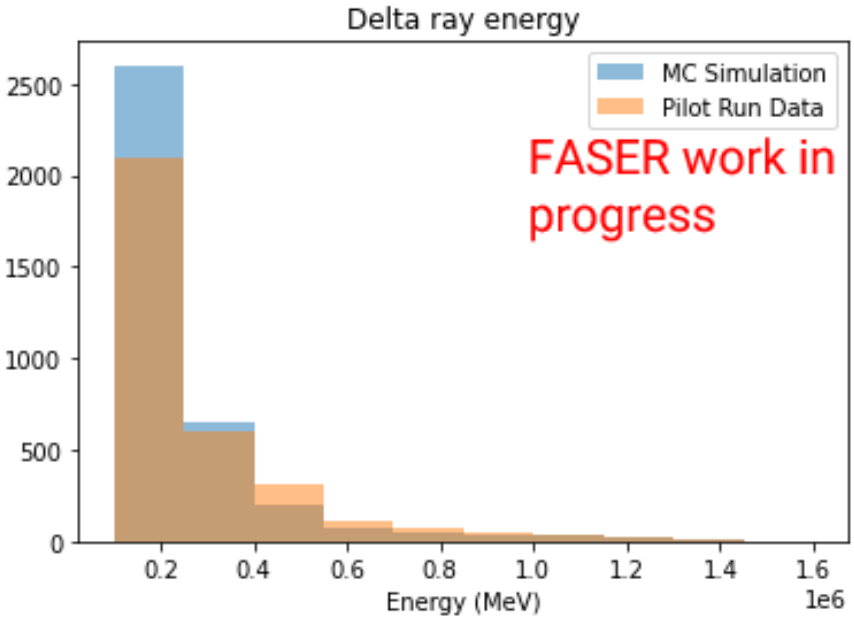

Figure 7.13: Overlapping histograms showing the comparison of reconstructed electron energies in the GEANT4 MC simulation (blue) and in the FASERnu pilot run data (orange).

actual data recorded by the detector. A small sample was taken from the lead module of the pilot detector, located ±2 mm from the position $(x, y) = (50, 50)$mm.

The expected number of ν_e events in the pilot detector is much less than 1, so they will not be considered. A much more interesting feature of the pilot run data is the presence of high-energy gammas/electrons; in principle it is possible to detect these $e^+/e^-/\gamma$ and measure their energy. Since there are very limited data in forward region, such data is unique and important to validate the simulations and demonstrate the energy measurement. One can first estimate the expected rate of $e^+/e^-/\gamma$ with $E > 500$ GeV (or $E > 1$ TeV) with the GEANT4 pilot detector simulation. All the other hits (muons, hadrons, remaining gammas/electrons) are background hits.

For ν_e events, there will be an electron starting inside the detector and there will be other tracks starting from the same emulsion plate converging with the impact parameter within a few microns. Having just a single electron without other hadrons/fragments is very rare.

For the ν_e study in the physics run, it is necessary to check neighbouring tracks for each vertex track. For the $\gamma/e^+/e^-$ study in the pilot run data, vertex reconstruction is not required, but rather a full check of the volume if there are tracks with many shower tracks around. If the energy is ¿500 GeV or so, the number of shower tracks would be much higher than the background hit level.

To perfrom the gamma/electron study in the pilot run data (also for ν_e energy measurement in the physics run simulation), the background hits must be implemented. Ideally this can be done with the simulation, but it is realistic to use data. The hits in the sub-volume provided can be used as background hits. The hit information (density, angles) can be added in simulation to check how the energy measurement will be affected. Thus a reasonable first step is to plot the angular distributions.

7.6 Comparison of clustered electron showers between MC and data

One of the most useful functions of the FASERν pilot detector is its ability to reconstruct electromagnetic showers arising from a high energy electromagnetic particle (i.e., e^+, e^-, or γ). In particular, it is important to reconstruct as many basetracks corresponding to daughter particles of the primary incident electron as possible. The motivations for this are

as follows:

- By counting the number of EM shower tracks, it is possible to reconstruct the energy of the incident primary electron. This can be used to estimate the energy of the neutrino which produces the electron in a CC event, and by counting the number of neutrinos produced in each energy bin, it is possible to measure the electron neutrino CC cross-section. Thus, electron EM shower reconstruction is important for the determination of the electron neutrino CC cross-section sensitivity.

- High-energy muons incident on the pilot detector are able to produce *delta rays* (also known as *knock-on electrons*), which are fundamentally electrons knocked off their host atoms by an incident particle. The reconstructed electron energy can be used to estimate the energy of the incident muon initiating the delta-ray, and by counting the number of muons in each energy bin, it is possible to obtain a reconstructed muon spectrum. Thus, electron energy reconstruction can be used to validate the simulated muon energy spectra.

The first task is to prepare an appropriate MC simulation sample to emulate the pilot run detector. It is known that the density of muons incident on the pilot detector is approximately $2.1 \times 10^5 (\mu^+ + \mu^-)/\text{cm}^2$, divided roughly evenly between μ^+ and μ^-. Using the energy and angular spectra, a realistic μ^+/μ^- sample can be generated with the pilot detector simulation. Since the pilot run data is described by the last 29 layers of the square cross-section $(50, 50)\,\text{mm} \oplus (12.5, 12.5)\,\text{mm}$, with cross-sectional area $6.25\,\text{cm}^2$, it is necessary to fire approximately $6.5 \times 10^5\,\mu^+/\text{cm}^2$ and $6.5 \times 10^5\,\mu^-/\text{cm}^2$ to emulate the pilot detector.

The placement of the pilot detector in the TI18 maintenance tunnel was performed by hand; this meant that there was no possibility for precision alignment, and as a result the pilot detector is not aligned with the beam axis, but rather offset so that its θ_x and θ_y distributions peak at $\theta_{\text{offset,x}} \approx -8 \times 10^{-3}\,\text{rad}$ and $\theta_{\text{offset},y} \approx 2.8 \times 10^{-2}\,\text{rad}$, as shown in Figure 7.14. Since the recorded data are actually $tx := \tan\theta_x$ and $ty := \tan\theta_y$, the alignment of the simulated

data is modified using the formula

$$tx_{\text{corrected}} = \tan[\tan^{-1}(tx_{\text{simulated}}) + \theta_{\text{offset},x}], \qquad ty_{\text{corrected}} = \tan[\tan^{-1}(ty_{\text{simulated}}) + \theta_{\text{offset},y}]. \tag{7.1}$$

After the angular corrections are applied, a comparison of the simulated and real data show

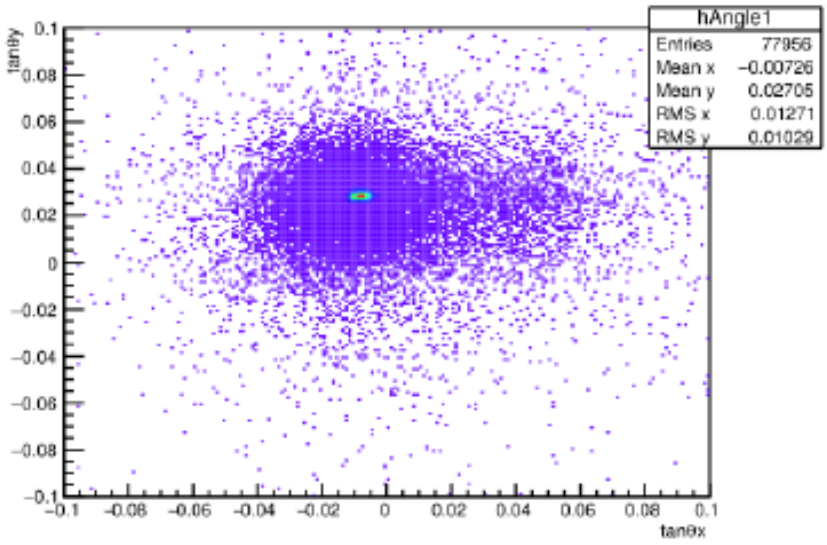

Figure 7.14: Scatter plot of actual angular spectra (θ_x, θ_y) of reconstructed basetracks in the pilot run.

that they match very precisely, as shown in Figure 7.15 It is interesting to investigate the source of the high-energy electrons measured in this study. While it is known that they are sourced by the parent muons, the claim that they are produced via muon-initiated delta-rays (as opposed to muon bremsstrahlung) must be verified.

Figure 7.16 shows a comparison of the transverse coordinates (x, y) of each reconstructed basetrack in the pilot run data and those of the virtual basetracks in the simulation. Note that there is an apparent deviation for x and y close to the boundaries of the pilot data sample; this is because the pilot sample is a subvolume extracted from the larger detector volume, whereas the simulation sample was generated within this subvolume only. As a

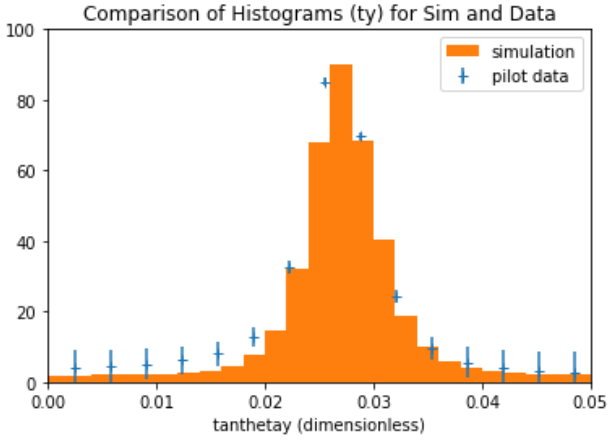

Figure 7.15: Comparison of histograms of $\tan\theta_x$ and $\tan\theta_y$ for simulation (solid color) and pilot data (error bars).

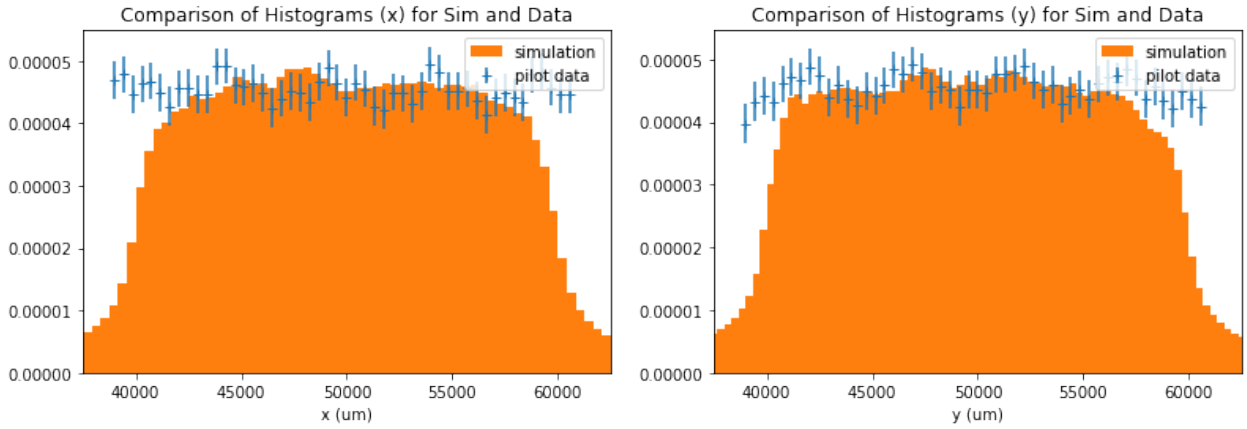

Figure 7.16: Comparison of histograms of x and y for simulation (solid color) and pilot data (error bars). The separation is due to spreading of simulation hits from inside the subvolume to outside, not balanced by spreading from outside the subvolume to inside (as with the pilot detector).

result, the basetracks inside the subvolume can spread out of the subvolume for both the simulation and the pilot run data; this effect is balanced by basetracks outside the subvolume spreading into the subvolume only for the pilot data.

Figure 7.17 shows the longitudinal plate index iz of the simulation sample compared

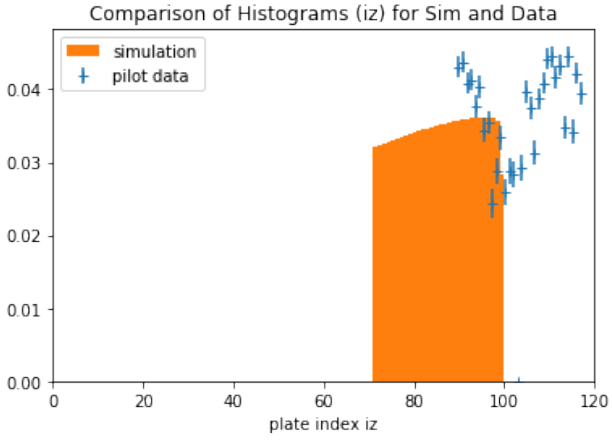

Figure 7.17: Comparison of histograms of iz for simulation (solid color) and pilot data (error bars). The inconsistency does not affect the data/MC comparison, see main text.

to the pilot data. Here 30 layers of the emulsion detector are simulated, compared to the 29 layers of pilot run data that have been provided. There is a difference in the above distribution because the simulation records the last 29 layers as plates 72 – 100, whereas the data sample records them as plates 90 – 118. This does not affect the data/MC comparison because after 90 emulsion plates the differences in energy and angular spectra are expected to be negligible.

One example high-energy electron event is shown in Figure 7.18. Note that the muon-initiated delta-ray is almost parallel to the direction of the incident muon (which scatters through a very small angle in the production of the delta-ray) and consequently goes directly through the middle of the electron EM shower.

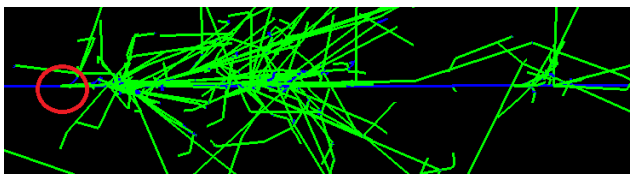

Figure 7.18: OpenGL GEANT4 event display of a muon (blue) which initiates a delta-ray (green) at a vertex (red).

7.7 Electron clustering

In order to reconstruct the knock-on electron energies, it is necessary to reconstruct the electron clusters. Figure 7.19 shows several truth knock-on electron clusters generated in

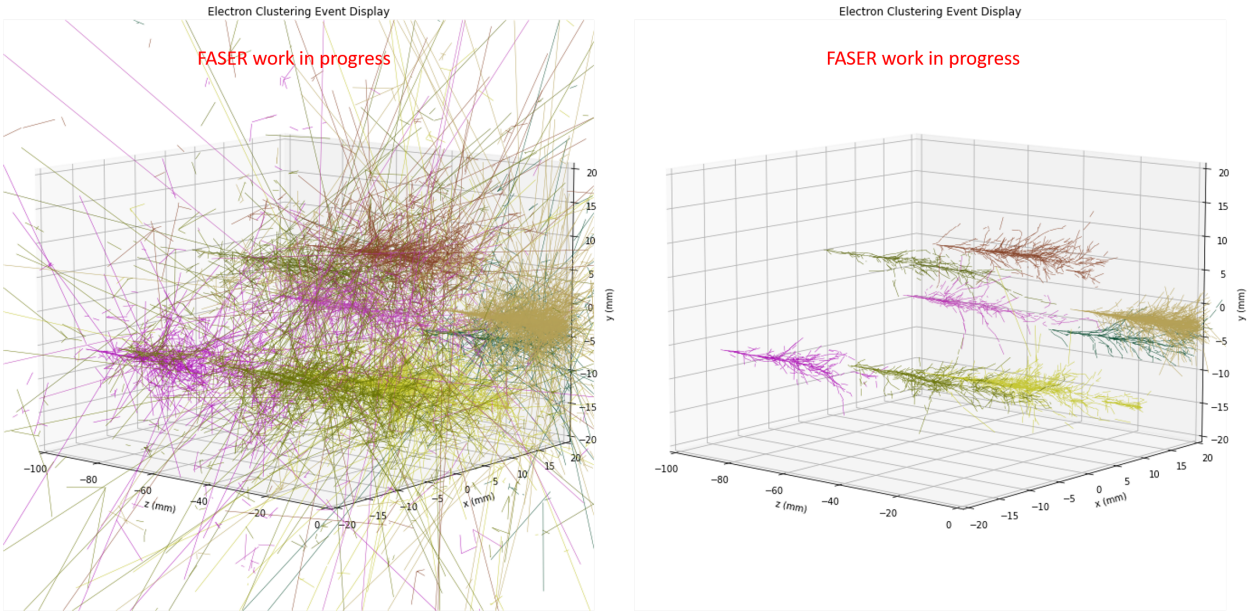

Figure 7.19: Sample of truth knock-on electron clusters (left) and reconstructed electron clusters (right) from MC sample.

the MC sample as compared to the electron clusters reconstructed from these truth clusters using the same seed basetracks. Note that not all truth cluster basetracks are reconstructed to the electron cluster.

In this study, the clustered basetracks (namely, those which clustering associates to this particular incident electron) are used, rather than the truth basetracks. The main advantage of this approach is that there is no "truth cluster" information in the pilot run data, but the clustering algorithm can be performed on simulation and data equally well.

Figure 7.20 shows a comparison of the number of reconstructed electron clusters in the MC simulation (blue) and in the pilot run data (orange) as a function of emulsion plate index (left) and energy (right). While there are more reconstructed electrons at lower energies in

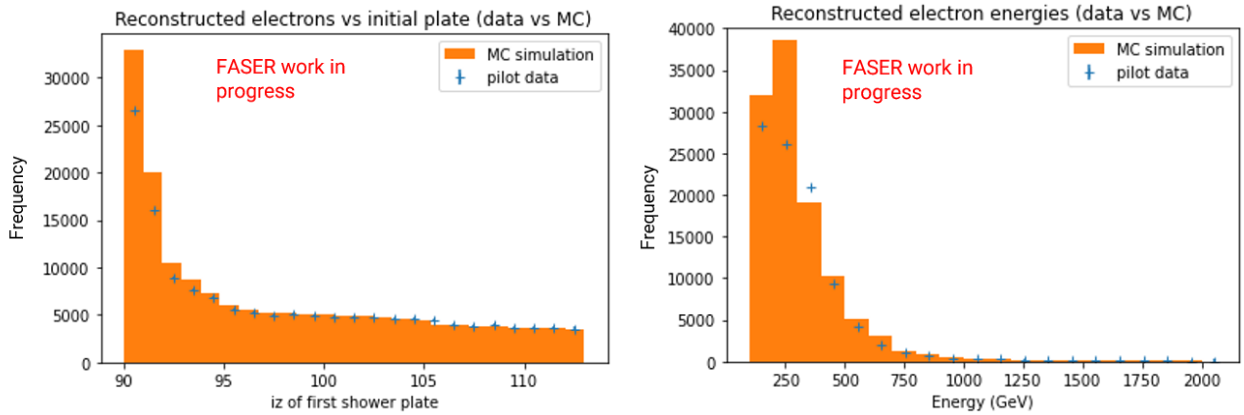

Figure 7.20: Number of clustered electron showers as a function of plate index iz (left) and of reconstructed electron energy (right).

the MC simulation, above $E = 250\,\text{GeV}$ there is excellent agreement.

Chapter 8

RESULTS AND CONCLUSION

The ForwArd Search ExpeRiment or FASER is an experiment designed to detect weakly interacting particles beyond the Standard Model. It is installed in the TI12 maintenance tunnel and has begun taking data as of April 2021. Particles register hits in SCT tracking stations with a dipole magnetic field in the volumes between the stations, and the track of a charged particle is reconstructed throughout the volume.

The FASERν detector is an emulsion detector to be placed upstream from FASER and to record the first neutrino interaction events at the LHC. It uses lepton identification to determine if a ν_e, ν_μ, or ν_τ has undergone a CC interaction. By reconstructing the energy of the interacting neutrino, the neutrino-nucleon CC cross-sections (averaged for neutrinos and antineutrinos) can be measured.

The FASERν pilot detector was installed in the TI18 maintenance tunnel (symmetric from TI12 with respect to the ATLAS interaction point) in 2018. It consists of a module with 120 0.5 mm thick tungsten absorber plates and another with 101 1 mm thick lead absorber plates, both interleaved with emulsion films. The pilot detector recorded $12.2\,\mathrm{fb}^{-1}$ of data during the pilot run from September 2018 to October 2018.

By combining the FASER and FASERν detectors using the IFT, it is possible to distinguish the muons produced by $\nu_\mu/\bar{\nu}_\mu$ interactions depending on their sign. The quality of the muon charge identification degrades with energy, but remains better than chance (i.e., the misidentification rate is less than 50%) even at energies as high as $E = 5\,\mathrm{TeV}$.

Electrons in FASERν and the pilot detector produce electromagnetic showers. Because the main source of these showers in the pilot detector is muons originating from the ATLAS interaction point, for which there are simulated energy and angular spectra, reconstructing

the energies of the EM showers is used to validate the simulation spectra. In addition, this procedure must be used to reconstruct the energy of a ν_e CC event, and this provides validation of the electron energy measurement approach.

Electron energy measurement is achieved by clustering together basetracks in the EM shower initiated by the electron. This cluster is then purified by removing basetracks with large impact parameters relative to the cluster seed, and the basetracks are counted to perform a one-dimensional electron energy regression. The electron energy resolution using only one-dimensional regression with the total number of clustered basetracks is already good enough to reconstruct electrons with energies of at least 300 GeV in the pilot detector.

Software used to produce the results in this thesis can be found at the git repository

`https://gitlab.cern.ch/jwspence/fasernu_pilot_analysis`

which contains detailed instructions for reproducing these results.

The author gave an APS talk [36] and DPF talk [37] regarding the measurement of neutrino/antineutrino CC cross sections with FASER and FASERν.

During the writing of this thesis, the author was a co-author of papers [33, 21, 24].

Chapter 9

OUTLOOK

The projects presented in this thesis so far suggest the following:

- The FASER spectrometer is a tracking device with good spatial and momentum resolution, and the spatial resolution improves with energy while the momentum resolution degrades with energy.

- The electron clustering algorithm is able to identify basetracks in the emulsion detector which belong to electron clusters with good purity.

- The one-dimensional electron energy regression algorithm using the number of clustered basetracks can reconstruct electrons with energies above 300 GeV with resolution better than $\sim 20\%$.

With tracker performance as good as the resolution suggests, it will be possible to reconstruct the parameters of charged particles – namely μ^+ and μ^- – to a high degree of precision. This enables the identification and distinction of μ^+ from μ^-.

Prospectively, the tracks in the FASER spectrometer and the FASERν emulsion detector can be matched to identify whether a straight track recorded in FASERν belonged to a μ^+ or μ^-. Together with the vertex identification for a ν_μ vertex, this enables the determination of whether a ν_μ or $\bar{\nu}_\mu$ underwent a CC interaction in FASERν. By counting the numbers of such events in each energy bin, it is then possible to make measurements of neutrino-nucleon CC cross-sections for ν_μ and $\bar{\nu}_\mu$ separately, rather than just the average of the two cross-sections.

In Chapters 6-7 the electron clustering and energy reconstruction algorithm was developoed and tested on simulation, then applied to the FASERν pilot data. While the full

algorithm can run on simulation and data alike, there are still differences which are not yet fully understood. This remains to be studied in future work.

The electron energy regression study has validated the algorithm for estimating electron energy by using the number of clustered basetracks associated to that electron. By using this same algorithm, one can reconstruct the energy of an electron emerging from a ν_e or $\bar{\nu}_e$ CC interaction, which will produce an electron or positron that immediately undergoes an EM shower.

The energy of the neutrino can then be estimated from the energy of the lepton through the process of unfolding. Note that a beam of neutrinos interacting at a fixed energy will produce a continuous spectrum of charged leptons as outgoing particles from the CC interaction. Therefore it is necessary to carry out an inversion task to estimate the neutrino energy from the outgoing lepton energy. (This is the case for both electron- and muon-type neutrinos.) This unfolding work is beyond the scope of this thesis, but is a necessary step to finalize the neutrino CC cross-section sensitivity measurement.

There are still ongoing studies with FASER and FASERν. Current efforts include a detailed study of ghost rejection (to remove fake basetracks that are reconstructed from the same particle in the same plate). Also, muon tracking is needed to identify the optimal criteria to be used to identify muons from emulsion tracks. In addition to the tracker performance in the full detailed simulation, charge misidentification requires further study. Furthermore, track matching between the FASER spectrometer and the FASERν emulsion detector is necessary to distinguish ν_μ from $\bar{\nu}_\mu$. Lastly, taking the charge misidentification rate and track matching as sources of systematic uncertainty, a revised $\nu_\mu/\bar{\nu}_\mu$ CC cross-section sensitivity would then be possible.

BIBLIOGRAPHY

[1] Akitaka Ariga et al. FASER: ForwArd Search ExpeRiment at the LHC. 1 2019.

[2] Henso Abreu et al. Detecting and Studying High-Energy Collider Neutrinos with FASER at the LHC. *Eur. Phys. J. C*, 80(1):61, 2020.

[3] C. Baltay et al. $\nu_\mu - \nu_e$ Universality in Charged Current Neutrino Interactions. *Phys. Rev. D*, 41:2653, 1990.

[4] M. G. Aartsen et al. First observation of PeV-energy neutrinos with IceCube. *Phys. Rev. Lett.*, 111:021103, 2013.

[5] K. Kodama et al. Final tau-neutrino results from the DONuT experiment. *Phys. Rev. D*, 78:052002, 2008.

[6] N. Agafonova et al. Latest results of the OPERA experiment on nu-tau appearance in the CNGS neutrino beam. *SciPost Phys. Proc.*, 1:028, 2019.

[7] Z. Li et al. Measurement of the tau neutrino cross section in atmospheric neutrino oscillations with Super-Kamiokande. *Phys. Rev. D*, 98(5):052006, 2018.

[8] S. F. Novaes. Standard model: An Introduction. In *10th Jorge Andre Swieca Summer School: Particle and Fields*, pages 5–102, 1 1999.

[9] M. Tanabashi et al. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

[10] G Rajasekaran. The Story of the Neutrino. 6 2016.

[11] F. Reines, C. L. Cowan, F. B. Harrison, A. D. McGuire, and H. W. Kruse. Detection of the free anti-neutrino. *Phys. Rev.*, 117:159–173, 1960.

[12] Ahmed Ismail, Roshan Mammen Abraham, and Felix Kling. Neutral Current Neutrino Interactions at FASERν. 12 2020.

[13] Petar Danev, Andrzej Adamczak, Dimitar Bakalov, Emiliano Mocchiutti, Mihail Stoilov, and Andrea Vacchi. Low-energy negative muon interaction with matter. *Journal of Instrumentation*, 11(03):P03019–P03019, Mar 2016.

[14] L. Chatterjee and B. K. Bandyopadhyay. Interaction Characteristics of Muons with Matter for the Energy Range 10**9-MeV >= e >= 10**7-MeV. *Acta Phys. Austriaca*, 45:345–357, 1976.

[15] Stephen Myers. The Large Hadron Collider 2008-2013. *International Journal of Modern Physics A*, 28(25):1330035–80, October 2013.

[16] J.D. Adam, Thierry Boutboul, G. Cavallari, Zinour Charifoulline, C.-H Denarie, Sandrine Le Naour, Df Leroy, L.R. Oberli, D. Richter, A.P. Verweij, and R. Wolf. Status of the lhc superconducting cable mass production. *Applied Superconductivity, IEEE Transactions on*, 12:1056 – 1062, 04 2002.

[17] Jason Nielsen. Fundamentals of LHC Experiments. In *Theoretical Advanced Study Institute in Elementary Particle Physics: String theory and its Applications: From meV to the Planck Scale*, pages 127–152, 6 2011.

[18] E. Lopez Sola et al. Beam impact tests of a prototype target for the Beam Dump Facility at CERN: experimental setup and preliminary analysis of the online results. *Phys. Rev. Accel. Beams*, 22(12):123001, 2019.

[19] Markus Cristinziani. The SHiP experiment at CERN. In *3rd World Summit on Exploring the Dark Side of the Universe*, 9 2020.

[20] Marco Fabbrichesi, Emidio Gabrielli, and Gaia Lanfranchi. The Dark Photon. 5 2020.

[21] Henso Abreu et al. The trigger and data acquisition system of the FASER experiment. 10 2021.

[22] Akitaka Ariga et al. Technical Proposal for FASER: ForwArd Search ExpeRiment at the LHC. 12 2018.

[23] Jamie Boyd. Forward search experiment at the lhc. RAL Seminar, 2020.

[24] Henso Abreu et al. The tracking detector of the FASER experiment. 12 2021.

[25] John Neil Jackson. The ATLAS semiconductor tracker (SCT). *Nucl. Instrum. Meth. A*, 541:89–95, 2005.

[26] Henso Abreu et al. Technical Proposal: FASERnu. 1 2020.

[27] Akitaka Ariga. Detecting and studying high-energy neutrinos with faserν at the lhc. ICHEP 2020, 2020.

[28] Stefan Roesler, Ralph Engel, and Johannes Ranft. The Monte Carlo event generator DPMJET-III. In *International Conference on Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications (MC 2000)*, pages 1033–1038, 12 2000.

[29] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner. EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider. *Phys. Rev. C*, 92(3):034906, 2015.

[30] S. Ostapchenko. QGSJET-II: Towards reliable description of very high energy hadronic interactions. *Nucl. Phys. B Proc. Suppl.*, 151:143–146, 2006.

[31] Felix Riehn, Ralph Engel, Anatoli Fedynitch, Thomas K. Gaisser, and Todor Stanev. Hadronic interaction model Sibyll 2.3d and extensive air showers. *Phys. Rev. D*, 102(6):063002, 2020.

[32] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[33] Henso Abreu et al. First neutrino interaction candidates at the LHC. 5 2021.

[34] W. Dabrowski. Readout of silicon strip detectors. *Nucl. Instrum. Meth. A*, 501:167–174, 2003.

[35] T. Kobayashi, Y. Komori, K. Yoshida, K. Yanagisawa, J. Nishimura, T. Yamagami, Y. Saito, N. Tateyama, T. Yuda, and R. J. Wilkes. OBSERVATIONS OF HIGH-ENERGY COSMIC-RAY ELECTRONS FROM 30 GeV TO 3 TeV WITH EMULSION CHAMBERS. *The Astrophysical Journal*, 760(2):146, nov 2012.

[36] John Spencer. Studying neutrino cross-sections with faser and fasernu. APS April Meeting, 2021.

[37] John Spencer. Measuring neutrino-nucleon cross-sections with fasernu. APS Division of Particles and Fields, 2021.

[38] Akitaka Ariga et al. Letter of Intent for FASER: ForwArd Search ExpeRiment at the LHC. 11 2018.

[39] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1–29, 2012.

[40] Katherine Freese. Status of Dark Matter in the Universe. *Int. J. Mod. Phys.*, 1(06):325–355, 2017.

[41] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helv. Phys. Acta*, 6:110–127, 1933.

[42] M. Eidemuller, Hans Gunter Dosch, and M. Jamin. The Field strength correlator from QCD sum rules. *Nucl. Phys. B Proc. Suppl.*, 86:421–425, 2000.

[43] V. Guzey and M. V. Polyakov. SU(3) systematization of baryons. 12 2005.

[44] W. Greiner and Berndt Muller. *Gauge theory of weak interactions*. 1993.

[45] A. Pich. The Standard model of electroweak interactions. In *2004 European School of High-Energy Physics*, pages 1–48, 2 2005.

[46] V. A. Bednyakov, N. D. Giokaris, and A. V. Bednyakov. On Higgs mass generation mechanism in the Standard Model. *Phys. Part. Nucl.*, 39:13–36, 2008.

[47] Máximo Bañados and Ignacio A. Reyes. A short review on Noether's theorems, gauge symmetries and boundary terms. *Int. J. Mod. Phys. D*, 25(10):1630021, 2016.

[48] Fulvio Laudisio. High precision reconstruction of electromagnetic showers in the nuclear emulsions of the opera experiment. page 640, 11 2017.

[49] Frédéric Juget. Electromagnetic shower reconstruction with emulsion films in the OPERA experiment. *Journal of Physics: Conference Series*, 160:012033, apr 2009.

[50] Xiaocong Ai. Acts: A common tracking software, 2019.

[51] Corey James Adams. *First Detection of Low Energy Electron Neutrinos in Liquid Argon Time Projection Chambers*. PhD thesis, Yale U., 2016.

[52] Francesco Spanò. Unfolding in particle physics: A window on solving inverse problems. *EPJ Web of Conferences*, 55:03002–, 07 2013.

Appendix A

HELIX TRACKER RECONSTRUCTION USING FASERMC

The FASER spectrometer is able to reconstruct the trajectories of charged particles. The presence of the uniform magnetic field in the dipole volumes but not the tracking stations implies that the particles will follow piecewise, smooth linear/helical tracks.

The GEANT4 simulation FaserMC is the legacy FASER simulation which predates the Calypso offline software framework. Note that this is a very simplistic model of the detector geometry and that Calypso was used for the more detailed simulation.

A.0.1 Naive Track Fitting

The first approach to tracker reconstruction was originally developed for the `FaserMC` legacy GEANT4 simulation. This simulation uses a simplified FASER detector geometry which includes only the SCT tracking stations and the calorimeter module.

After running the GEANT4 simulation, the naive track fitting algorithm is performed by performing a χ^2 fitting of a hybrid track as shown in Figure A.1. This is an improvement to the legacy fitting code, which performed a χ^2 fitting of a pure helix track.

The hybrid track-fitting is performed by performing a χ^2 fitting of a track that is a helix in the dipole volumes and a straight line within the tracking stations. The piecewise track is continuous and differentiable at all points. The helical sections are also constrained to have the same radius due to the uniform magnetic field. The detector geometries with and without the IFT are considered; the geometry without the IFT leads to three linear segments and two helical segments, whereas the geometry leads to an additional segment of each type for each charged track.

The fitting model for the helix uses five fitting parameters: the radius R, the initial

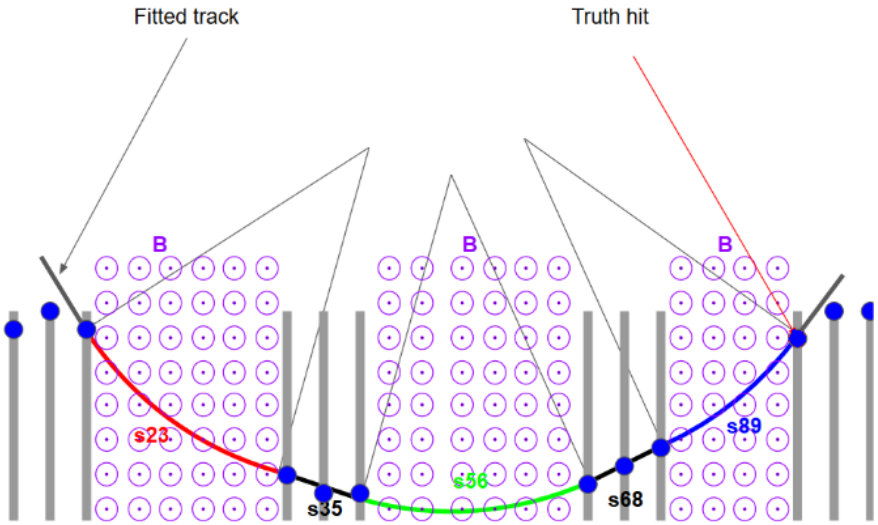

Figure A.1: A schematic diagram showing the hits registered by a μ^+ in the FASER spectrometer as well as the hybrid track that the particle takes while travelling through it. The tracking stations are shown in grey, the hits are shown in blue, and the direction of the nonzero magnetic field in the dipole volumes between the tracking stations is shown in magenta. The black track segments are linear while the coloured segments are helical.

transverse position (x_0, y_0), the initial azimuthal angle $\phi_=0$, and the helix pitch λ. The designations of these fitting parameters are shown in Figure A.2. If t denotes the arc length traveled along the helix, then the transverse arc length is $t\cos\lambda$ and the transverse angle is $(t/R)\cos\lambda$. Thus the fitted helix obeys the equations

$$x(t) = x_0 + R\left[\cos\left(\frac{t}{R}\cos\lambda + \phi_0\right) - \cos\phi_0\right], \tag{A.1}$$

$$y(t) = y_0 + R\left[\sin\left(\frac{t}{R}\cos\lambda + \phi_0\right) - \sin\phi_0\right], \tag{A.2}$$

$$z(t) = z_0 + t\sin\lambda, \tag{A.3}$$

$$\tag{A.4}$$

While it may appear that the new hybrid fitting model has more parameters due to the separate helices, this is merely an illusion. There are indeed multiple helix track segments, but the radii and pitches of all such helix segments are equal due to the magnetic field; in addition, the condition that the linear segments of the track connect to the helical arcs

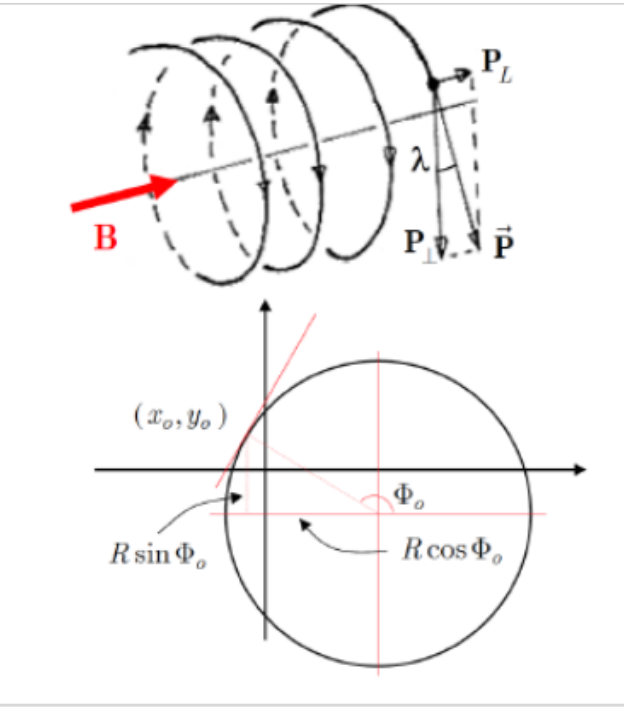

Figure A.2: Illustration of a helix and its projection to the plane transverse to its axis, along with the track radius R, the initial transverse position (x_0, y_0), the initial azimuthal angle ϕ_0, and the helix pitch λ.

continuously and differentiably imply that x_0, y_0, and ϕ_0 are fixed for all helix segments in terms of their values for the first helix segment, as shown in figure A.1. In effect, there are still only five free parameters.

One important question is how much the resolution improves using the realistic hybrid-fitting model in the inhomogeneous magnetic field over the legacy helix-fitting model in the uniform magnetic field. Another is how much the resolution improves by adding the interface detector as a fourth tracking station. One would expect no significant improvement with the modified fitting model and no IFT, and moderate improvement when the IFT is included.

The performance study using the new hybrid tracker was conducted as follows:

1. 10^5 μ^+ and 10^5 μ^- were fired at the FASER detector using the legacy `FaserMC` GEANT4 simulation tool twice, once with a homogeneous 0.55 T magnetic field everywhere and

once with the modified inhomogeneous field that is nonzero only in the dipole volumes outside the tracking stations.

2. The GEANT4 hits were fitted to a helix track using the helix tracker ROOT macro for the simulation with the uniform magnetic field.

3. The GEANT4 hits were fitted to a hybrid helix/linear track using the hybrid tracker ROOT macro for the simulation with the inhomogeneous magnetic field.

4. The parameters of the reconstructed helix (hybrid) tracks and the truth helix (hybrid) tracks were compared and plots of the resolutions of each of the five parameters were generated.

The results of this performance study were as follows.

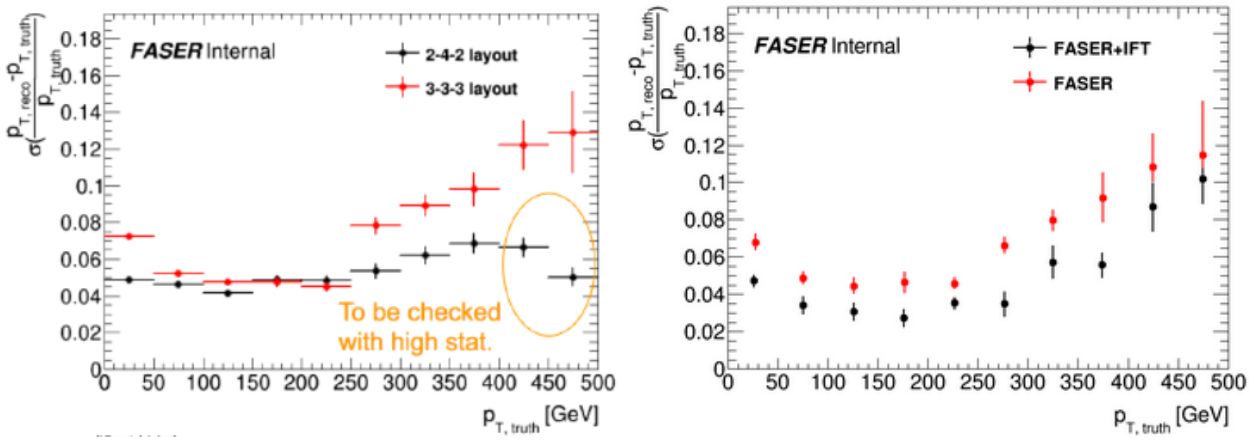

Figure A.3: (left) p_T resolution of FASER spectrometer using the helix tracker and the legacy MC simulation; (right) p_T resolution of FASER spectrometer using the improved hybrid tracker.

Figure A.3 shows the momentum resolution of the FASER spectrometer. The p_T resolution when using a uniform magnetic field and helix tracks is shown on the left, whereas the p_T resolution using the true magnetic field and the hybrid helix/linear tracks described in

Figure A.1 above is shown on the right. Evidently, using the hybrid tracks improves the p_T resolution slightly, while adding the IFT improves resolution significantly.

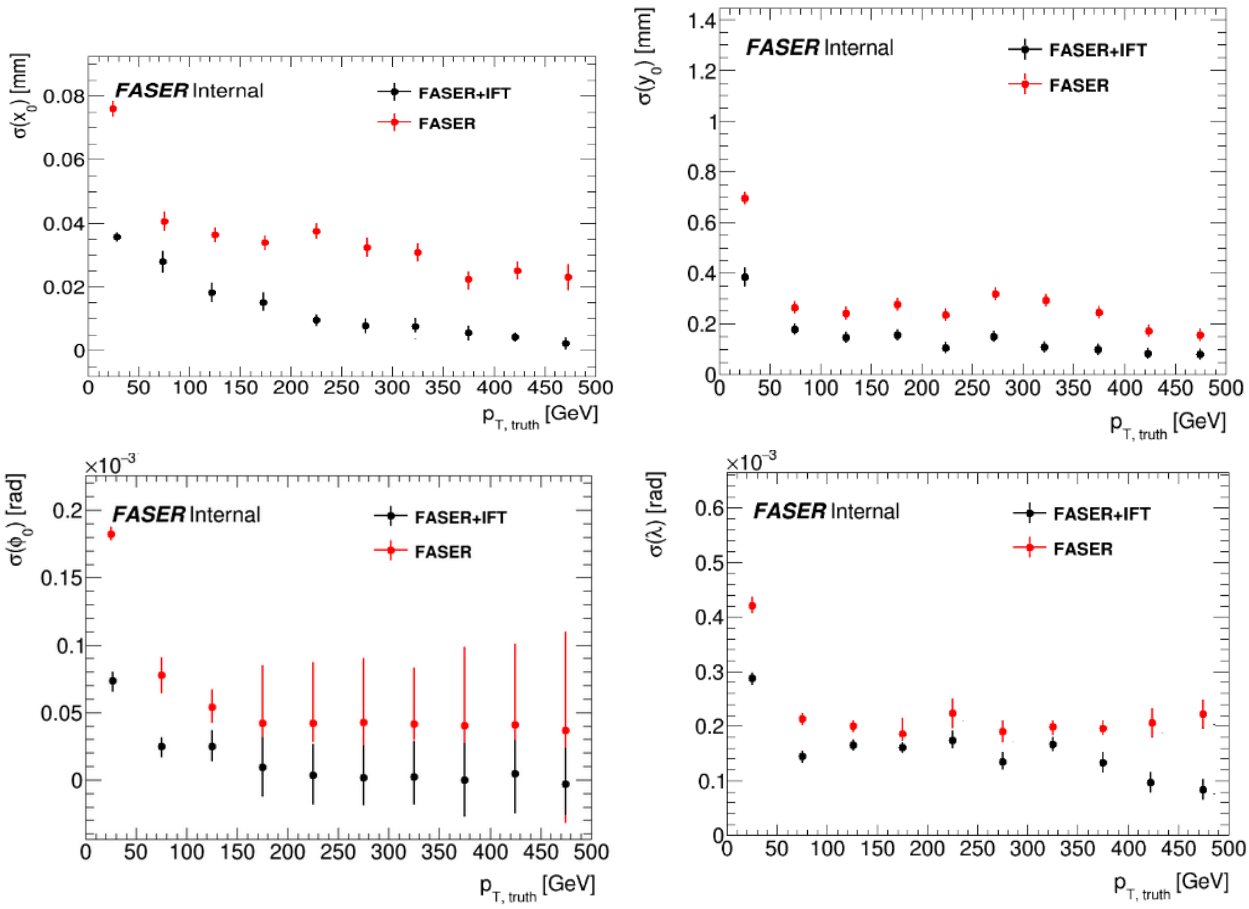

Figure A.4: Resolutions for the helix fitting parameters (top-left) x_0, (top-right) y_0, (bottom-left) ϕ_0, and (bottom-right) λ using the hybrid helix/linear tracker and the legacy MC simulation. Data points indicate resolution for tracks reconstructed without the IFT (red) and with the IFT (black).

Figure A.4 demonstrates the tracker performance without the IFT (red) and with the IFT (black). As is evident, the resolution for each of the helix-fitting parameters improves when the IFT is introduced.

Appendix B

NEUTRINO SIGNAL / BACKGROUND SIGNATURES

Now that the theory background has been introduced and the experimental apparatus has been described, the signatures of the signal processes (ν_e, ν_μ, and ν_τ charged-current events) and the background processes are overviewed. The separation of signal from background using characteristics of those hits and their tracks. This is necessary to identify different flavours of neutrino events.

B.1 Electron neutrino CC events

As discussed in Chapter 1, the neutrinos can undergo charged-current events producing charged leptons. Under the assumption that lepton number is conserved, the lepton number of the neutrino (i.e., whether it is a neutrino or an antineutrino) is anticorrelated with the charge of the outgoing lepton. In particular, for electron neutrino CC events, an incoming ν_e will produce an outgoing e^-, while an incoming $\bar{\nu}_e$ will produce an outgoing e^+.

Figure B.1 illustrates these three neutrino-nucleon CC interactions in the emulsion detector.[2] The vertical lines indicate the emulsion layers interleaving the tungsten absorbers. These three signatures are identified in [26], and the emulsion detector will provide the functionality necessary to distinguish between them. In what follows, we focus on the signature of the ν_e CC event.

When an electron strikes an absorber layer in the detector, it undergoes bremsstrahlung, losing a fraction of its energy to a photon. When a photon strikes an absorber layer, it can produce an $e+/e-$ pair. These electrons will also undergo bremsstrahlung and the process will repeat until the electron energy falls below the critical energy, at which points collisions become the dominant form of energy loss. This process is known as an *electromagnetic*

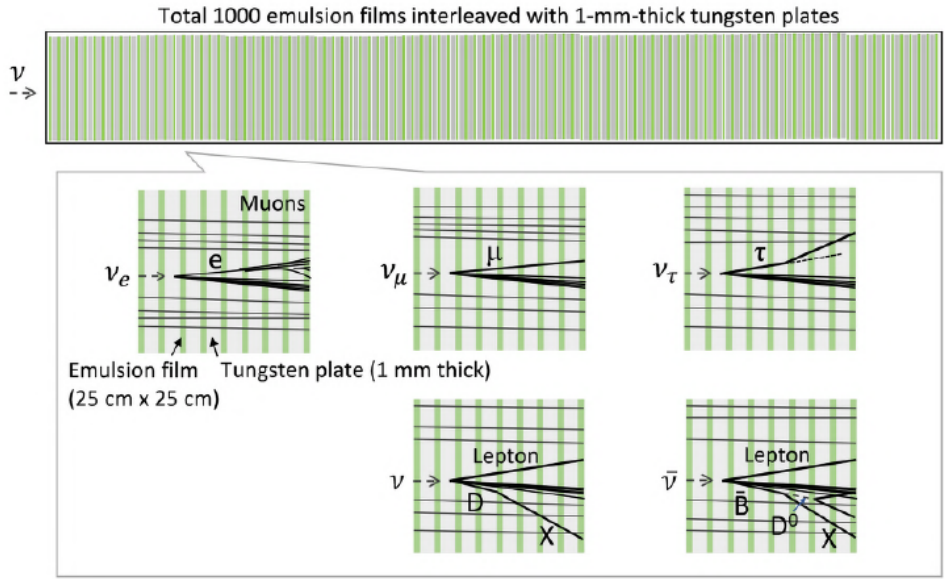

Figure B.1: Lepton signatures for each of the three neutrino-nucleon CC interactions.

shower.

It is immediately apparent that an energetic electron incident on the FASERν detector will initiate an EM shower in the first layer of the FASERν detector. The energy of the electron decreases exponentially with depth until it falls below the critical energy E_c. The expected distance an electron travels in a material before losing all but $1/e$ of its energy due to bremsstrahlung, pair-production and other processes is known as the *radiation length* of the material and is denoted by X_0; for tungsten, the radiation length is $X_0 \approx 3.504$ mm, whereas for lead, $X_0 \approx 5.612$ mm. Thus, the signature of an electron neutrino CC event is a single, large EM shower. As will be demonstrated, the typical depth of the electron-induced EM shower is approximately 100 layers.

Figure B.2 illustrates an example electromagnetic shower. Note that this diagram has been simplified to remove the quark components necessary for an accurate depiction of the pair-production process. In the above figure, the incident electron undergoes bremsstrahlung, emitting a photon which interacts with the detector to produce an electron-positron pair.

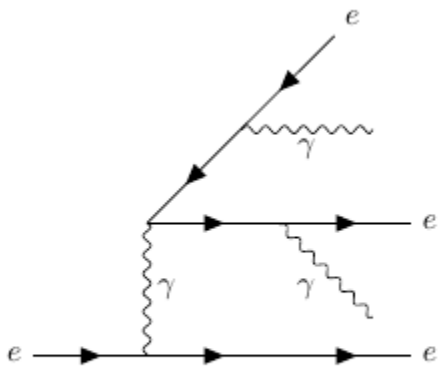

Figure B.2: Feynman diagram for a simplified electromagnetic shower. The incident electron emits a photon, which splits into an electron-positron pair, each of which emit further photons.

The daughter electron and positron each undergo further bremsstrahlung, emitting photons.

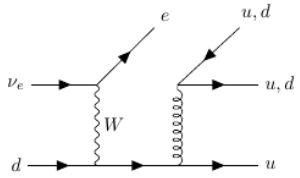

Figure B.3: Feynman diagram of electron neutrino CC event, which initiates a hadronic shower by stimulating emission of a virtual gluon which decays into a neutral pion.

In addition to the electromagnetic shower, high-energy electrons may also initiate hadronic showers. As shown in Figure B.3, hadronic showers are initiated e.g. by stimulated nucleon pion-production. These pions may interact strongly with subsequent layers of the emulsion

detector to stimulate emission of more pions, and so on until there is insufficient energy for further pion production.

In analogy to the definition of the radiation length X_0 in terms of energy loss via electromagnetic processes, the *interaction length,* denoted L_0, is defined as the length over which a hadron loses all but $1/e$ of its energy through hadronic processes. The interaction length in tungsten for pions is approximately 113.3 mm; by comparison to the interaction length in tungsten of 3.504 mm, it is clear that any ν_e-induced EM shower in the FASERν emulsion detector would occur significantly upstream of any hadronic shower initiated by the same process.

B.2 Muon neutrino CC events

Incoming muon neutrinos undergoing charged-current interactions will produce outgoing muons. Unlike the outgoing electrons from ν_e CC events, the outgoing muon from a ν_μ CC events does not initiate a single large EM shower; in contrast, the increased mass of the muon prevents it from losing a significant fraction of its momentum through a single bremsstrahlung event. As a result, the outgoing photon carries a significantly lower energy. This lower energy corresponds to a smaller EM shower. Furthermore, the high outgoing momentum of the muon means it is susceptible to further bremsstrahlung. Therefore, instead of a single large EM shower, one expects multiple small EM showers along the muon track.

In addition to this behavior, the mu neutrino can also stimulate pion production in exactly the same way as the electron neutrino. Indeed, the hadronic showers resulting from a ν_μ CC event are so much larger than the many small EM showers from the same event. Therefore, the signature of a ν_μ CC event is a hadronic shower with a number of smaller EM showers and an emerging μ which reaches the end of FASERν.

B.3 Tau neutrino CC events

Unlike the electron neutrino and muon neutrino CC events, the tau neutrino CC event produces an unstable τ particle which promptly decays in the emulsion detector. Given the

lifetime of the τ is 2.903×10^{-13} s, the mean expected flight distance of a τ in FASERν is ~ 3 cm. Therefore, since τ particles produced in ν_τ CC events are expected to be focused in the forward direction, it is expected that any τ thusly produced will cross at least one emulsion layer before decaying.

The most common τ decay modes are $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$, $\tau^- \to e^- \bar{\nu}_e \nu_\tau$, and $\tau^- \to \pi^- \nu_\tau$ (mutatis mutandis for τ^+). In all cases, the outgoing lepton is invisible, but the charged particle (μ^-, e^-, or π^-) will record hits in the emulsion detector. These are known as *1-prong decays* because only one visible particle is produced. A tau neutrino will produce an outgoing τ that will subsequently undergo a 1-prong decay approximately 85% of the time. It is therefore expected that the signature of a ν_τ CC event is the production of a charged track with a "kink" or directional discontinuity which indicates the decay of the τ.

The identification of ν_τ CC events is performed by identifying a track segment that crosses at least one emulsion layer, and further requiring that the kink angle is greater than both 0.5 mrad and four times the angular resolution of FASERν given the length of the particle. Additionally, it is required that the track of the reconstructed τ be less than 6 cm in length to reduce hadronic background; as will be discussed in the next section, the hadronic interaction length in tungsten is approximately 10 cm, so it is expected that hadronic tracks will exceed this interaction length.

. This procedure is estimated to perform correct identification of ν_τ CC events approximately 75% of the time [2]. It is also possible for a τ to undergo a 3-prong decay. The most common of these decay modes is the decay to three charged pions and a neutrino, $\tau \to \pi^- \pi^- \pi^+ \nu_\tau$, which occurs with branching fraction 8.9868 ± 0.0513 [9]. In this case the identification is performed by identifying the three secondary tracks, and the τ ID efficiency is estimated to be at least 75% if not higher.

B.4 Background Processes

In addition to the above neutrino CC events, there are background processes which must be considered. The first of these backgrounds is due to the muons incident on FASERν. The

expected energy spectra of incident μ^+ and μ^- is shown in Figure B.4. These muons will also initiate hadronic showers just as the signal ν_μ CC events do; however, there is a large flux of muons expected from outside of LHC and so it becomes necessary to separate the background showers from those in the signal events.

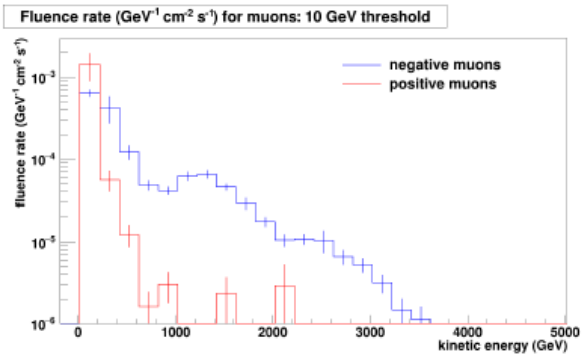

Figure B.4: Expected distribution of background positive (red) and negative (blue) muons incident on FASERν.

These incident muons contribute significantly to backgrounds in two ways:

1. Muons traveling through FASERν can initiate multiple small EM showers along their trajectory.

2. Muons traveling through the rock upstream from FASERν will create secondary e/γ .

When the data from FASERν are analyzed, there will be many hits due to both signal and background. In order to reconstruct neutrino events, it will be necessary to distinguish the hits corresponding to a signal event of a neutrino-induced electron initiating an EM shower from the hits corresponding to one of the main backgrounds outlined above. Then it will be possible to use only the hits that were identified as signal hits to reconstruct electron energy.

In order to establish the expected background, a `GEANT4` simulation of the rock upstream from FASERν was performed. This simulation included a trapezoidal geometry for the rock

and a single layer of FASERν to record the emerging spectra. The expected numbers of

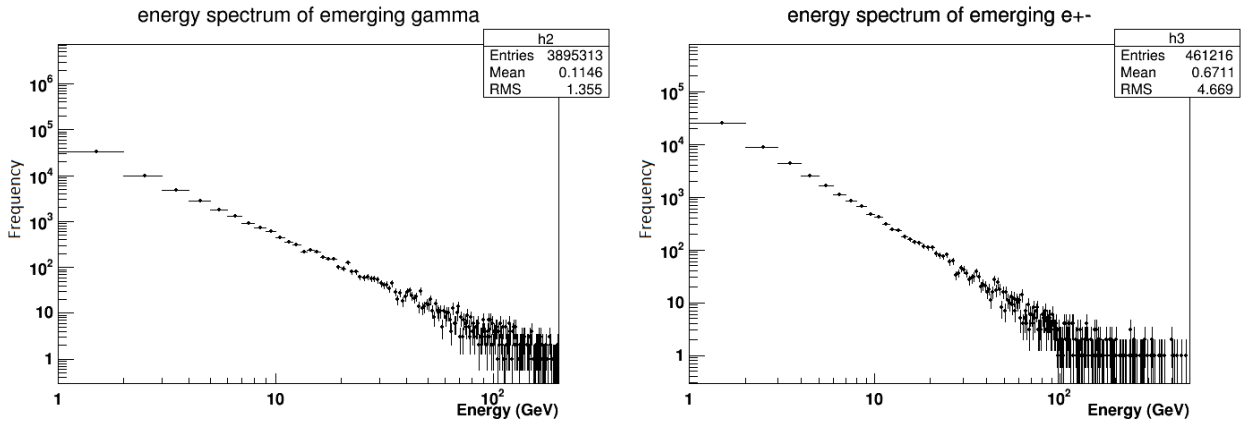

Figure B.5: Energy spectra for emerging photons (left) and e^+/e^- (right) emerging from incident μ^-.

background hadrons and photons, as well as the expected number of signal ν_μ and $\bar{\nu}_\mu$ CC events, is shown in Table B.8. It is expected that 2×10^9 muons, roughly evenly distributed between μ^+ and μ^-, will be incident on FASERν during Run 3 of the LHC.

Figures B.5- B.7 show the energy spectra of the most relevant backgrounds resulting from muons incident on the rock upstream from FASERν with the spectra given in Figure B.4. All of these emerging secondary particles are significant for the following reasons:

- Emerging $e^+/e^-/\gamma$ will initiate EM showers when they come into contact with FASERν. These can be conflated with the EM showers initiated by e^+/e^- emerging from a ν_e CC event.

- Neutral hadrons will not be detected by the emulsion detector but initiate hadronic showers, which can be mistaken for the signature of a ν_μ CC event.

- Muons will initiate smaller EM showers which will mask the e^+/e^- hits corresponding to ν_e CC events.

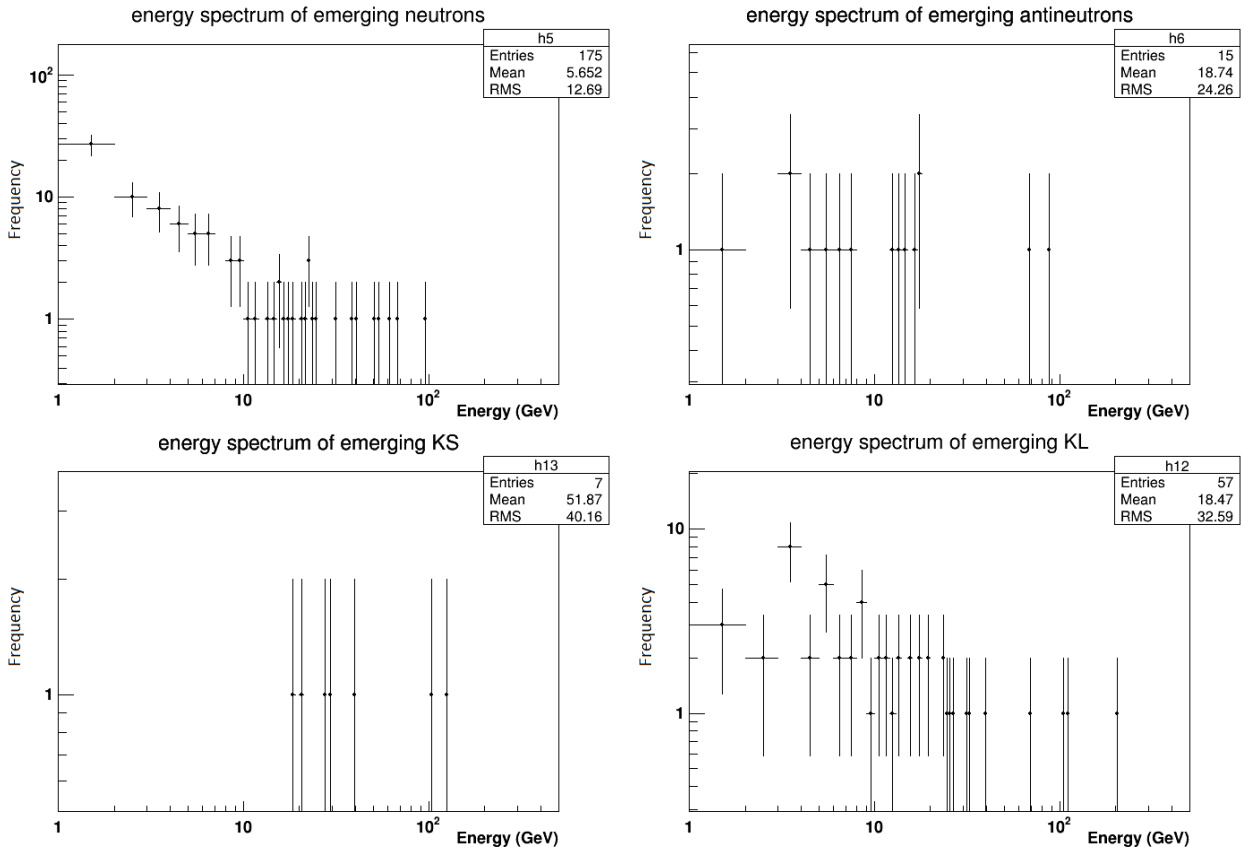

Figure B.6: Energy spectra for emerging neutrons (top-left), antineutrons (top-right), K_S (bottom-left), and K_L (bottom-right) emerging from incident μ^-.

Note that the multiple EM showers initiated by muons will result in an approximately uniform hit distribution along the length of FASERν, whereas the single large EM shower initiated by electrons will create a peaked distribution in the first ~ 100 emulsion layers, as shown in Figure B.9.

It is important to distinguish the e^+/e^- hits from the signal ν_e CC events and the e^+/e^- hits due to the background μ background events. The major characteristics that assist in this distinction are the energies, directions, and transverse distance of the hits. The hits from a signal event are high-energy and collimated in the forward direction, whereas the background

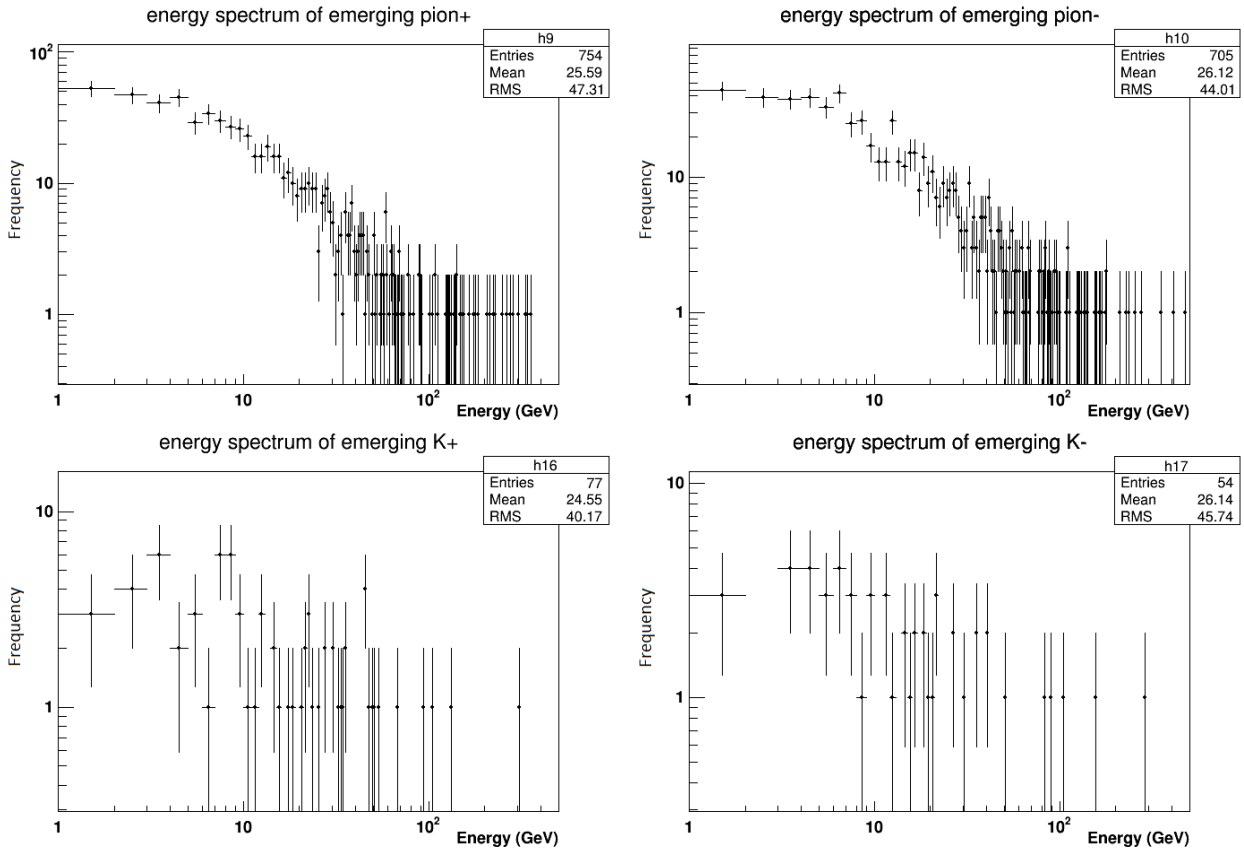

Figure B.7: Energy spectra for emerging π^+ (top-left), and π^- (top-right), K^+ (bottom-left), and K^- (bottom-right) emerging from incident μ^-.

hits are lower-energy and more spread out. By applying appropriate cuts, namely $E > 100$ MeV, $\tan\theta < 0.05$, and $x^2 + y^2 < (200\,\mu\text{m})^2$, it is possible to reduce the background to a negligible contribution, as shown in Figure B.10.

Particle	Expected number of particles passing through FASERν			
	$E > 10$ GeV	$E > 100$ GeV	$E > 300$ GeV	$E > 1$ TeV
Neutrons n	27.8k / 138k	1.5k / 11.5k	150 / 1.1k	2.2 / 42
Anti-neutrons $\bar{n}$	15.5k / 98k	900 / 9k	110 / 1.5k	2.8 / 46
Λ	5.3k / 36k	390 / 4.1k	39 / 800	0.9 / 58
Anti-Λ	3.4k / 31k	290 / 3.5k	31 / 200	0.6 / 14
K_S^0	1.3k / 30k	240 / 6.8k	52 / 390	1.8 / 6.2
K_L^0	1.6k / 31k	270 / 5.7k	55 / 500	1.2 / 18
Ξ^0	240 / 1.3k	13 / 190	2.3 / 12	0.1 / –
Anti-Ξ^0	150 / 1k	10 / 200	1.4 / 19	–
Photons γ	2.2M / 62M	160k / 16.3M	38.2k / 6.3M	5.9k / 1.1M
$\nu_\mu + \bar{\nu}_\mu$ (signal int.)	23.1k	20.4k	13.3k	3.4k

Figure B.8: Table of expected background induced by muons and signal particles passing through FASERν.

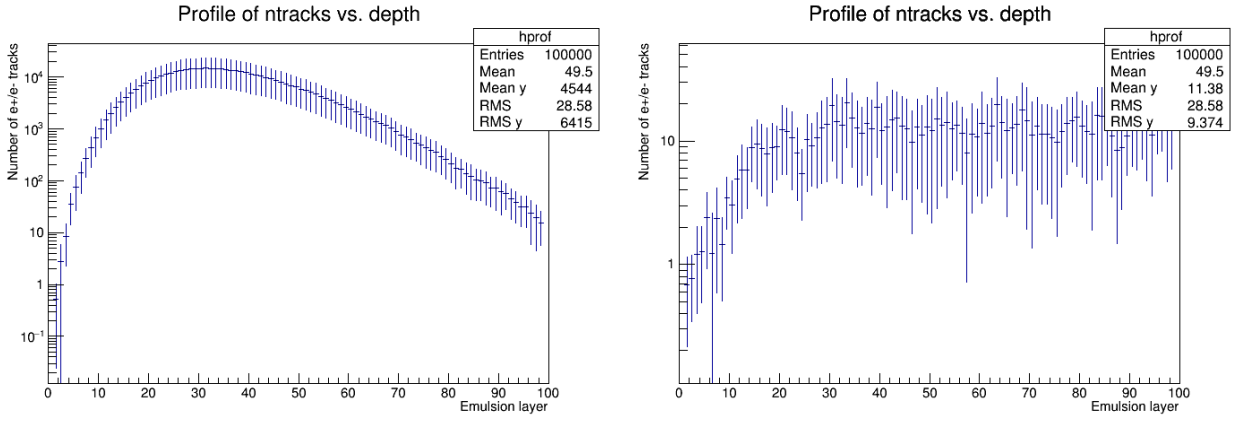

Figure B.9: Depth profile for 10^4 electrons (left) and negative muons (right) at 1 TeV energy and at normal incidence on FASERν.

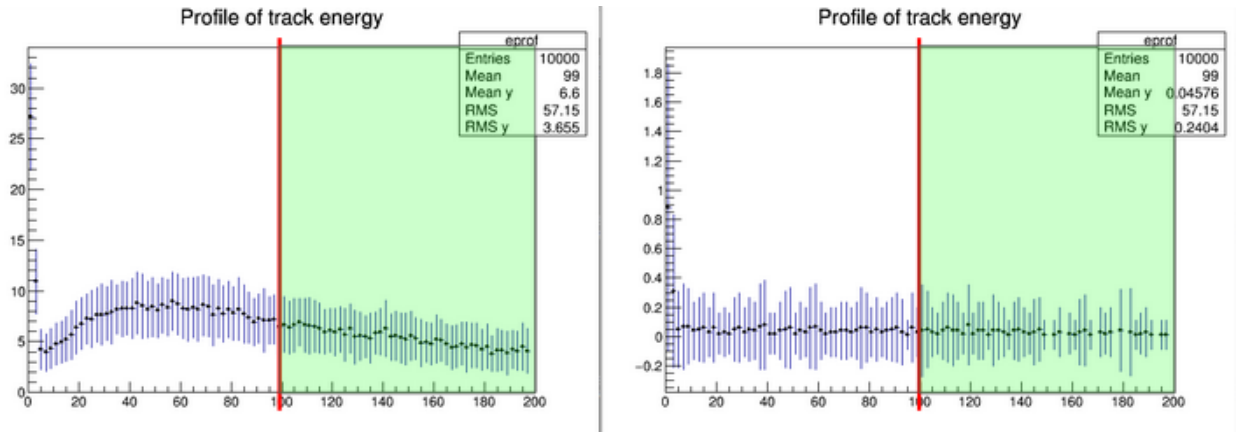

Figure B.10: Signal-background separation for 10^4 electrons (left) and negative muons (right) at 1 TeV energy with cuts $E > 100$ MeV, $\tan\theta < 0.05$, and $x^2 + y^2 < (200\,\mu\text{m})^2$.

Appendix C

TRUTH-LEVE SIMULATION STUDY OF ELECTROMAGNETIC SHOWERS

In this appendix the characteristics of electromagnetic showers are presented. The total number of electron tracks, the depth of the shower maximum, and the maximum number of tracks in a given layer are discussed, and the boosted decision tree is used to perform electron energy regression. The general behaviour of the electromagnetic shower depth profiles is introduced, and the artificial neural network is used as an energy regression technique.

C.1 Characteristics of Electromagnetic Showers

As discussed in Chapter 5, the signature of an electron neutrino CC event in FASERν is a large EM shower. Furthermore, the characteristics of an EM shower are dependent on the energy of the incident particle. Thus, using one or more variables associated to the EM shower, it should be possible to reconstruct the energy of the outgoing electron from the ν_e CC event (and thence the incident neutrino energy).

The variables which can be used for electron energy reconstruction are

1. The total number N_{total} of e^+/e^- tracks reconstructed in FASERν;

2. The maximum track multiplicity N_{max} of e^+/e^- tracks in a single layer;

3. The layer index iz_{max} in FASERν at which the track multiplicity is a maximum;

4. The track multiplicity N_ℓ in each layer of FASERν.

In practice, it is simpler to use the layer index iz_{max} rather than the physical depth z_{max} corresponding to this index. The behaviour of N_{total} and N_{max} is linear with the incident

electron energy E, while the behaviour of $z_{\max}$ (hence $iz_{\max}$) is logarithmic. The explicit formulae for $N_{\max}$ and $z_{\max}$ [48]:

$$z_{\max} = \frac{X_0}{\ln 2} \ln \frac{E_{\text{incident}}}{E_c}, \tag{C.1}$$

$$N_{\max} = \frac{E_{\text{incident}}}{E_c}. \tag{C.2}$$

The distribution of track multiplicity as a function of depth is also known as an *EM shower (depth) profile.*

C.2 Electron Energy Reconstruction with Total Track Count

A first approach is simply to use the total number of base tracks. Samples were generated using GEANT4 by firing muons into the rock simulation to estimate the spectra of emerging particles (e.g., $e+/e-/\gamma$). These output energy/angular spectra were then used as input spectra to the GEANT4 simulation of FASERν. 10^4 electrons were fired at FASERν at normal incidence and the resulting EM shower depth profiles were studied. The `ROOT` data analysis framework was used to count the total number of tracks and discard all tracks that were not e^+/e^- tracks. Figure C.1 shows the total number of tracks reconstructed in

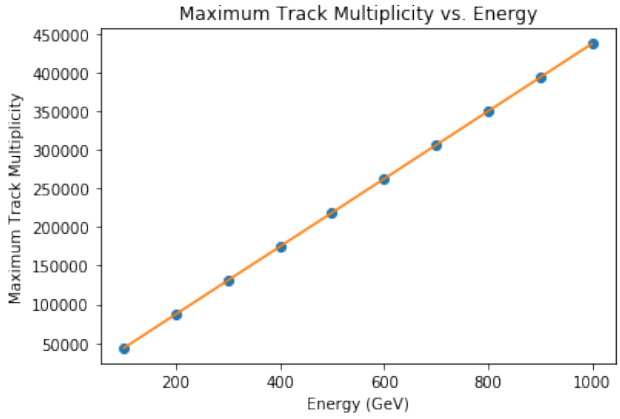

Figure C.1: Plots of the total number of tracks as a function of electron energy along with the fitted linear regression.

FASERν as a function of energy. A linear regression was fitted to the profile data:

$$\chi^2 := \frac{(N_{\text{total}}^{(\text{fitted})} - N_{\text{total}}^{\text{truth}})^2}{\sigma(N_{\text{total}})^2} \tag{C.3}$$

The simulated results for N_{total} as a function of energy is shown in Figure [REF]. There is adequate resolution for electron energy reconstruction. However, there is a significant drawback in that there is maximum uncertainty in the number of reconstructed incident electron – there is, for example, no distinction between one incident 100 GeV electron and two 50 GeV electrons. Therefore, electron energy reconstruction using only the total e^+/e^- track count is not a reliable method in practice.

C.3 Electron Energy Reconstruction with Maximum Track Depth and Multiplicity

The second approach for reconstructing incident electron energy is to use the additional variables of the shower maximum depth iz_{max} and the multiplicity N_{max} in that layer.

$$N_{\text{total}}^{(\text{fitted})}(E) = 437.0621 \left(\frac{E}{1\,\text{GeV}}\right) - 14.1450667, \tag{C.4}$$

$$N_{\text{max}}^{(\text{fitted})}(E) = 11.1921 \left(\frac{E}{1\,\text{GeV}}\right) + 330.0005, \tag{C.5}$$

$$iz_{\text{max}}^{(\text{fitted})}(E) = 3.854 \ln\left(\frac{E}{1\,\text{GeV}}\right) + 11.398. \tag{C.6}$$

Figure C.2 shows a linear regression of the profile of maximum track multiplicity, and a logarithmic regression of the profile of shower maximum depth, as functions of energy. It is apparent that the

This method of electron energy reconstruction can be realized with χ^2 fitting by minimizing the variable

$$\chi^2 := \frac{(N_{\text{total}}^{(\text{fitted})} - N_{\text{total}}^{\text{truth}})^2}{\sigma(N_{\text{total}})^2} + \frac{(N_{\text{max}}^{(\text{fitted})} - N_{\text{max}}^{\text{truth}})^2}{\sigma(N_{\text{max}})^2} + \frac{(iz_{\text{max}}^{(\text{fitted})} - iz_{\text{max}}^{\text{truth}})^2}{\sigma(iz_{\text{max}})^2}, \tag{C.7}$$

where $N_{\text{total}}^{(\text{fitted})}$ (resp. $N_{\text{total}}^{(\text{truth})}$) denote the fitted (resp. truth) total e^+/e^- track count, $N_{\text{max}}^{(\text{fitted})}$ (resp. $N_{\text{max}}^{(\text{truth})}$) the fitted (resp. truth) maximum track multiplicities, and $iz_{\text{max}}^{(\text{fitted})}$ (resp.

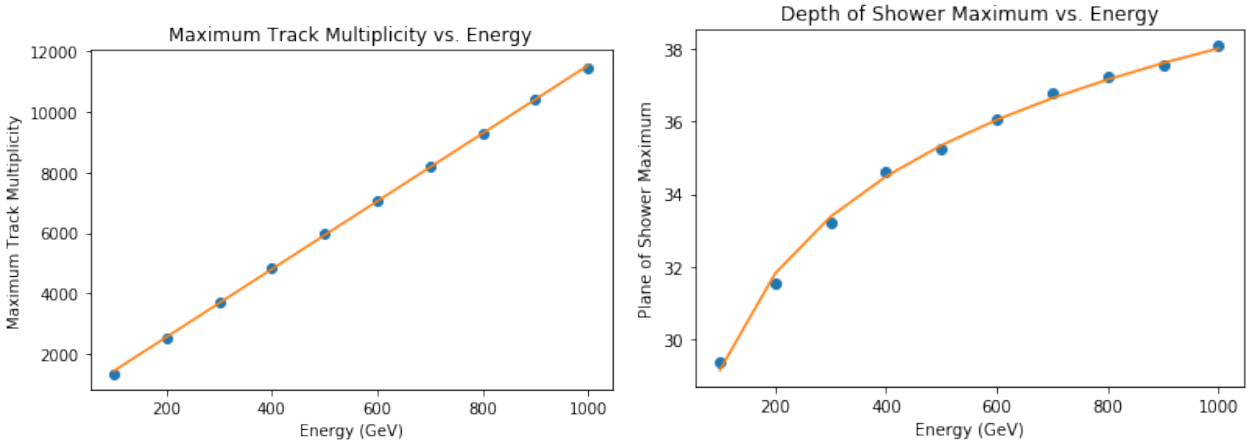

Figure C.2: Plots of the maximum track multiplicity in a single layer (left), and the depth of the shower maximum (right), as a function of electron energy. A linear regression is fitted to the track multiplicity, while a logarithmic regression is fitted to the depth of the shower maximum.

$iz_{\max}^{(\text{truth})}$) the fitted (resp. truth) depths of shower maxima. The energy for which $N_{\text{total}}, N_{\max}$, and $iz_{\max}$ minimizes χ^2 is the reconstructed energy.

The fitted equations for $N_{\text{total}}, N_{\max}$, and $iz_{\max}$ as functions of incident electron energy are given by

Performing a χ^2 fit of the observed values of these variables to the above equations gives the desired result.

Another realization of this electron energy reconstruction approach is to use *boosted decision trees*. This is a supervised machine learning technique that involves providing training data samples and constructing decision trees by making cuts on each of the variables $N_{\text{total}}, N_{\max}$, and $iz_{\max}$, branching repeatedly, and returning reconstructed values when a terminal node or leaf of the tree is reached. Boosted decision trees are reliable for regression and classification but are susceptible to overfitting, which reduces their predictive power. For this reason it is important to specify a maximum depth. In this case a maximum depth of 10 was initially used.

This approach has an improved energy resolution to that discussed in the previous section.

Indeed, there is now a distinction between one incident 100 GeV electron and two 50 GeV electrons in that the depth of the shower maximum is different between the two scenarios. In addition, it is possible to eliminate background muons by using the fact that the muon shower profiles in FASERν are effectively uniform due to the multiple small EM showers initiated along the muon track. This is a feature that the total track count approach lacked. However, this electron energy reconstruction approach also fails because the superposition of two distinct profiles may have $N_{\max}$ and $iz_{\max}$ which cannot be derived from $N_{\max}$ and $iz_{\max}$ for each individual profile.

C.4 Electron Energy Reconstruction with the Full Shower Profile

The final approach for electron energy reconstruction is to use the full EM shower depth profile. In this case, the track multiplicity at each of 100 layers is compared to the track multiplicities for each of those layers in the reference energy profiles.

This energy reconstruction approach has the best energy resolution of the three methods, and can also separate a single signal event from background. However, more importantly, it allows for linear superposition; it is possible, for example, to reconstruct multiple incident electrons by a brute-force search of the reference profiles. This is also the best possible energy reconstruction method available to FASERν because it uses all of the information recorded by the emulsion films.

As before

After generating the samples, we will reconstruct the tracks and study profiles of the variables E_{track}, $\tan\theta$, and $r = \sqrt{x^2 + y^2}$ where r is the transverse direction relative to the primary hit.

In addition to the background EM showers, the signal events may have more than one EM shower due to the following effects:

1. Neutrinos incident on FASERν can produce δ-rays (knock-on electrons) which will initiate EM showers.

2. Neutrinos incident on FASERν can lead to neutrino-nucleon π^0-production. The π^0 rapidly decays electromagnetically to photons which initiate further EM showers.

When we analyze the data recorded by FASERν or the pilot detector, we will not have just the hits from a single event but in fact hits from multiple ν_e events. These events will overlap and therefore the number of hits in each layer (used in Section 1 to reconstruct the electron energy) will not fit well to the profiles we use for energy reconstruction. Instead, it will be a superposition of such ntracks vs. depth distributions, which we will have to study how to separate efficiently.

In addition, when a ν_e interacts inside FASERν, we may get multiple EM showers resulting from the processes described above. In this case there will be a more complicated superposition of EM showers because the depths at which the showers are initiated will vary.

Samples will be generated using GEANT4. We will fire electrons at multiple energies from different positions at normal incidence. We will combine the hits of these events into a single event.

After the samples are generated, we will reconstruct the tracks. The number of tracks per layer will be recorded and we will attempt to separate them into the constituent EM showers, then perform electron energy reconstruction as in Section 1 for each of the identified electrons.

One possible, though computationally expensive, way to perform this pileup profile energy reconstruction is as follows:

1. Attempt to fit one EM shower profile at a reference energy (10 cases); record χ^2.

2. Attempt to fit the superposition of two EM shower profiles, each at a reference energy (100 cases); record χ^2 and use it instead if χ^2 is lower than the above.

3. Repeat until χ^2 is minimized, then use this as the reconstructed number of electrons n_E and the estimated energies $E_i, 1 < i < n_E$.

4. For each reconstructed electron, use χ^2 fitting with the logarithmic profile to reconstruct the energy as in Section 1.

A single incident electron neutrino can initiate multiple EM showers by the target nucleus producing a π^0, which then decays electromagnetically [51]. In this case, there would be hits from electrons that were members of multiple EM showers, which would be initiated both at different transverse positions (x, y) and at different depths z. This is a source of background that will be considered later.

The electron energy regression will be done using reference profiles using one of three methods:

1. Compare the *total* number of electron tracks in FASERν from the test event with the total number of electron tracks at each of the reference energies;

2. Compare the total number of electron tracks, the layer with the highest number of electron tracks, and this maximum number of electron tracks in one layer, for the test event and these three variables for each of the reference energies;

3. Compare the number of electron tracks in each layer from the test event with the numbers of electron tracks in each layer at each of the reference energies.

Intuitively, the last of these three approaches should have the best resolution, although this has yet to be studied. While initially the approach for energies between the reference energies was to interpolate linearly between the fixed values, a more statistically sensible approach is to perform regressions. In the case of the total and maximum numbers of electron tracks, this is accomplished through the usage of *boosted decision trees (BDTs).* This is a machine learning technique often used in both regression and classification tasks, and where the regression resolution generally improves with the depth of the BDT. BDTs are implemented using `python` with the `scikit-learn` package.

While the one- and three-dimensional electron energy regression algorithms are implemented with BDTs, this approach fails to be computationally efficient for the layer-by-layer energy profile reconstruction. Therefore the final energy regression technique is implemented using artificial neural networks (ANN), and in particular the `TMultiLayerPerceptron` package in `ROOT`. This ANN approach was used in [26] for muon neutrino energy reconstruction.

Appendix D

TRACKER AND TRACKER+IFT PERFORMANCE

As discussed previously, the tracking stations of FASER are separated by large dipole magnets of magnetic field 0.55 T. This magnetic field is sufficient to curve the paths of charged particles entering FASER into noticeable piecewise helical/linear trajectories. The handedness of the helical segments correponds exactly to the charge of the particle – a positively-charged particle will trace a right-handed helix in the dipole volumes between tracking stations, while a negatively-charged particle will trace a left-handed helix. Therefore, if we can precisely reconstruct these tracks, we can use this information to reconstruct the charge of the particle.

The reconstructed charge is determined from the sign of the q/p fitted track parameter. The sign of this parameter is compared to the truth charge to determine if the charge was misidentified. Note that the tracking stations lie entirely outside the magnetic field, and so the track seeds left in the tracking stations are indeed straight, as expected since the parameters $(x_0, y_0, \theta, \phi, q/p)$ do not explicitly determine a charged track. This requires a nonzero magnetic field B, at which point the radius r is determined by $1/r = (q/p)B$. Because of the evolution of the state vector as the particle traverses each dipole magnet, the track parameters describing the helices are different for the helical segments in each dipole volume.

The FASERν detector in isolation can distinguish between the three flavours of neutrino-nucleon CC interactions, but cannot distinguish between the interactions of neutrinos and antineutrinos; as such, FASERν in isolation can determine only the average neutrino/antineutrino CC cross section and not the CC cross section for neutrinos and antineutrinos separately. However, with the interface detector, the piecewise tracks fitted to a charged particle in the

FASER spectrometer can be matched to the linear emulsion tracks it records in FASERν. In conjunction with the property that Standard Model neutrinos and antineutrinos undergo CC interactions which produce charged leptons of different electric charge, one may perform separate measurements of the neutrino and antineutrino cross sections.

The charge reconstruction can be performed in one of three ways. The first approach is to fit a piecewise helical/linear track to the hits in all tracking stations, and then use the handedness of the helix to determine the sign of the charge – i.e., a right-handed helix indicates a charge of +1. A second approach is to make the charge itself one of the fitted variables (or a function thereof). The third approach is to use the sign of the *sagitta,* or perpendicular distance from the arc of the helix to its chord, to determine the sign of the charge. These three approaches were studied independently and were shown to have comparable results. Ultimately the second approach was selected – the fitted track parameters include q/p, the fitted value of the charge divided by the magnitude of the momentum.

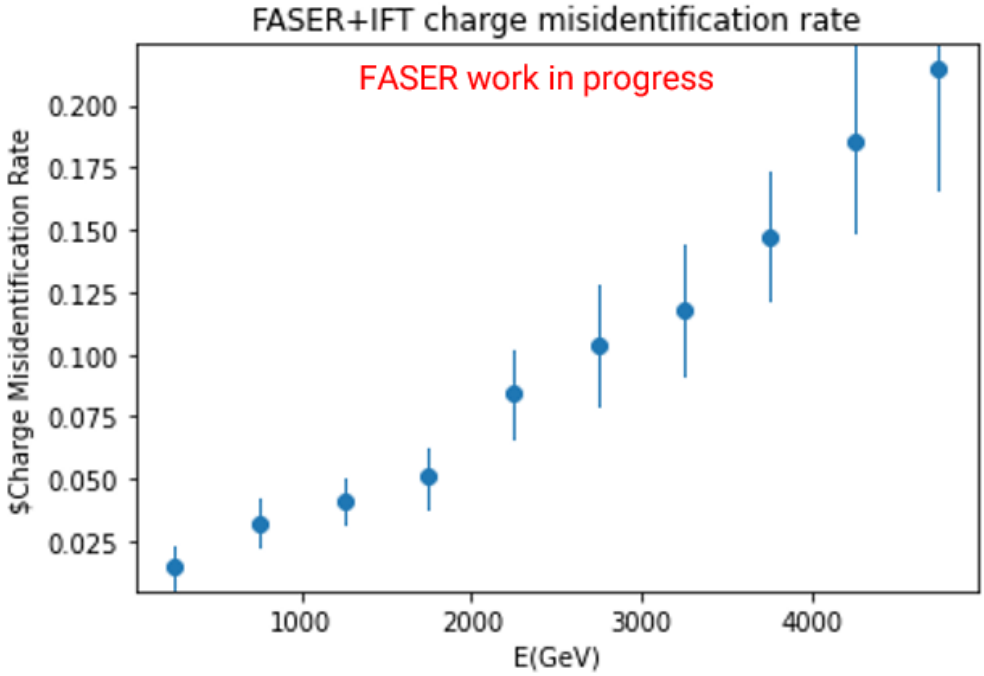

Figure D.1: Plot of the charge misidentification rate as a function of incident muon energy.

Figure D.1 shows the proportion of muons with misidentified charge as a function of their incident energy. The behaviour is as expected: At lower energies, the chirality of the helix is more clear and charge misidentification is unlikely, but as energy increases the helix better approximates a line and the frequency of misidentification increases. Observe that a misidentification rate of 0.5 is equivalent to random assignment of charge and that we should not expect charge misidentification to exceed 0.5 as a result.

Note that this charge misidentification rate is dependent on the muon energy, whereas what will be measured is the muon neutrino energy. For a given energy of incident muon neutrino energy, there is an energy spectrum for the produced muon. This spectrum needs to be taken into account when determining the charge misidentification rate as a function of neutrino energy; this process is known as *unfolding* [52]. Figure [REF] shows the unfolded muon charge misidentification rate. As expected, the misidentification rates are lower for the same energies; this is due to the fact that the outgoing muons have lower energy than the incident neutrinos, and therefore leave better-defined helix tracks in FASER.

In practice, the quantities needed to reconstruct the neutrino-nucleon CC cross-section are the numbers of reconstructed incident neutrinos and antineutrinos in the FASER detector. According to the standard model, these are the same as the true number of emerging μ^- and μ^+, respectively, and thus will be denoted $N_{\pm}^{\text{truth}}$ concordantly. If $N_{\pm}^{\text{reco}}$ denote the number of $\mu^{\pm}$ reconstructed in FASER that correspond to a muon neutrino CC event in FASERν, and $x(E)$ denotes the misidentification rate, then

$$\begin{cases} N_{+}^{\text{truth}} = (1 - x(E))N_{+}^{\text{reco}} + x(E)N_{-}^{\text{reco}}, \\ N_{-}^{\text{truth}} = x(E)N_{+}^{\text{reco}} + (1 - x(E))N_{-}^{\text{reco}}. \end{cases} \tag{D.1}$$

Therefore, once the number of reconstructed μ^+/μ^- tracks and their reconstructed energy is known, Equation D.1 can be used to obtain the expected number of μ^+/μ^- that actually interacted with FASER as a function of energy.

Because the determination of incident neutrino lepton number is contingent upon the resulting charged lepton reaching the end of FASERν and hitting the tracking stations of the FASER spectrometer, it is apparent that separate neutrino and antineutrino cross-section

measurements are possible for only ν_μ CC events (the electron will produce an EM shower, and the τ will decay, well before reaching the end of FASERν). The procedure is as follows:

1. A ν_μ CC event is detected in FASERν and the charged emulsion tracks reconstructed.

2. The emerging μ^+/μ^- reaches the end of FASERν, enters the FASER spectrometer, and leaves piecewise helical/linear tracks there.

3. The handedness of the reconstructed helical tracks determines the sign of the charge.

4. The interface detector is used to match the spectrometer tracks in FASER with the emulsion tracks in FASERν.

5. The μ emulsion track is traced back to the reconstructed ν_μ CC vertex.

6. The reconstructed muon charge is used to determine whether a ν_μ or $\bar{\nu}_\mu$ interacted in FASERν.

This functionality is clearly only possible through the interfacing of the FASER spectrometer with the FASERν emulsion detector, and that FASER can perform charge identification only because of the 0.55 T magnetic field that permeates it. Therefore it is of great importance how often this charge will be misidentified – that is, a μ^- with positive reconstructed charge or vice versa.

D.1 Tracker performance

After running the MC simulation samples through the full chain of digitization and CKF, two files are produced, containing truth and reconstructed track information for each sample. It is then possible to carry out a tracker performance study as follows:

1. Make plots of the residuals of reconstructed track parameters $(x, y, \theta, \phi, q/p)$ as a function of truth energy.

2. In each energy bin, the residuals form a one-dimensional histogram. Fit a Gaussian distribution to this histogram.

3. The mean of the fitted distribution is the *bias* and expresses whether there is a tendency of the algorithm to reconstruct the parameter systematically high or systematically low.

4. The standard deviation of the fitted distribution is the *resolution* and expresses the uncertainty associated with each measurement.

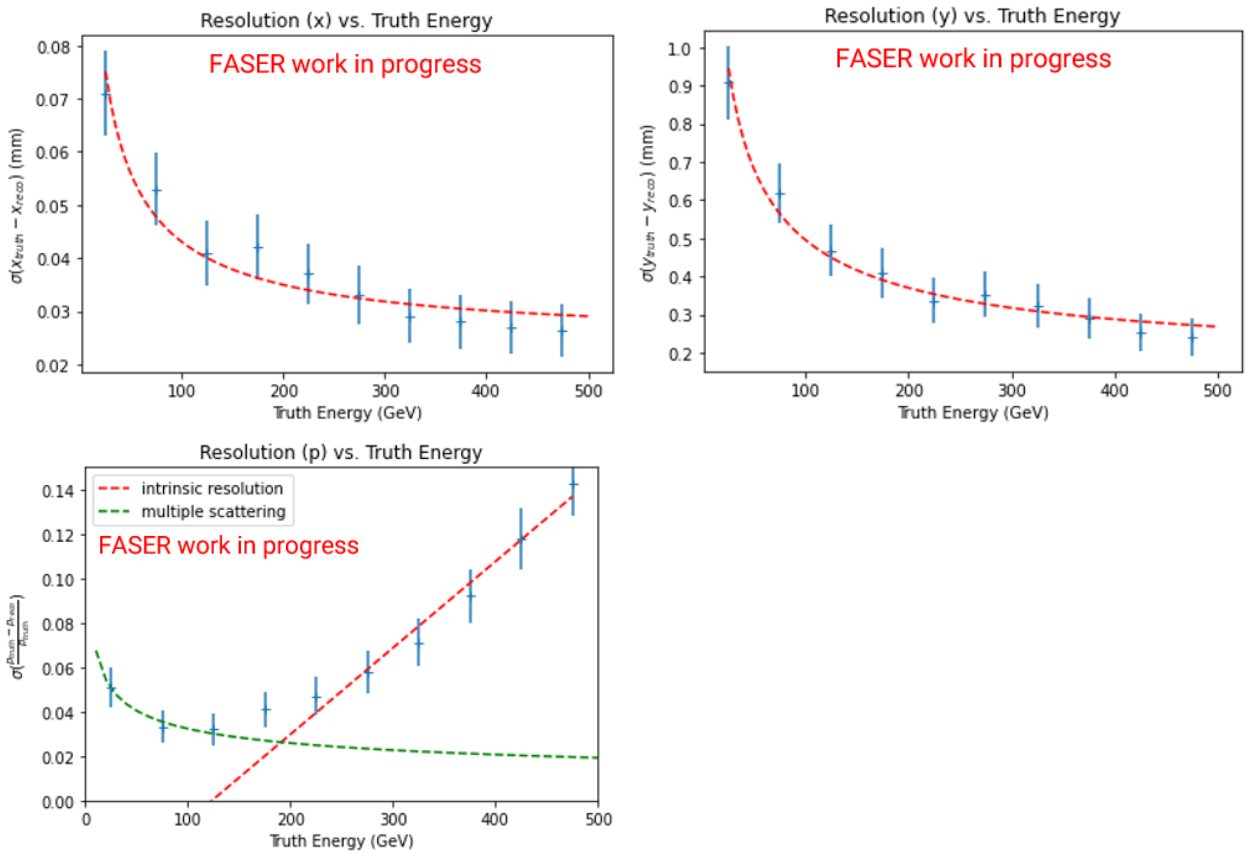

Figure D.2: Tracker x position resolution (top-left), y position resolution (top-right) and fractional momentum resolution (bottom-left) versus truth energy, with fitted curves.

Figure D.2 shows the tracker performance as a function of truth muon energy. It is evident that the spatial resolution improves with energy as the muons follow straighter paths in

space, but a consequence of this same fact is that the momentum resolution degrades with energy.

Appendix E

IFT TRACK MATCHING BETWEEN FASER TRACKER AND FASERν EMULSION DETECTOR

Figure E.1 shows a CAD model of part of the FASER spectrometer and the FASERν emulsion detector.[2] The distance between the last emulsion plate and the first module of the IFT is approximately 121.08 mm and is represented by the green line in the figure.

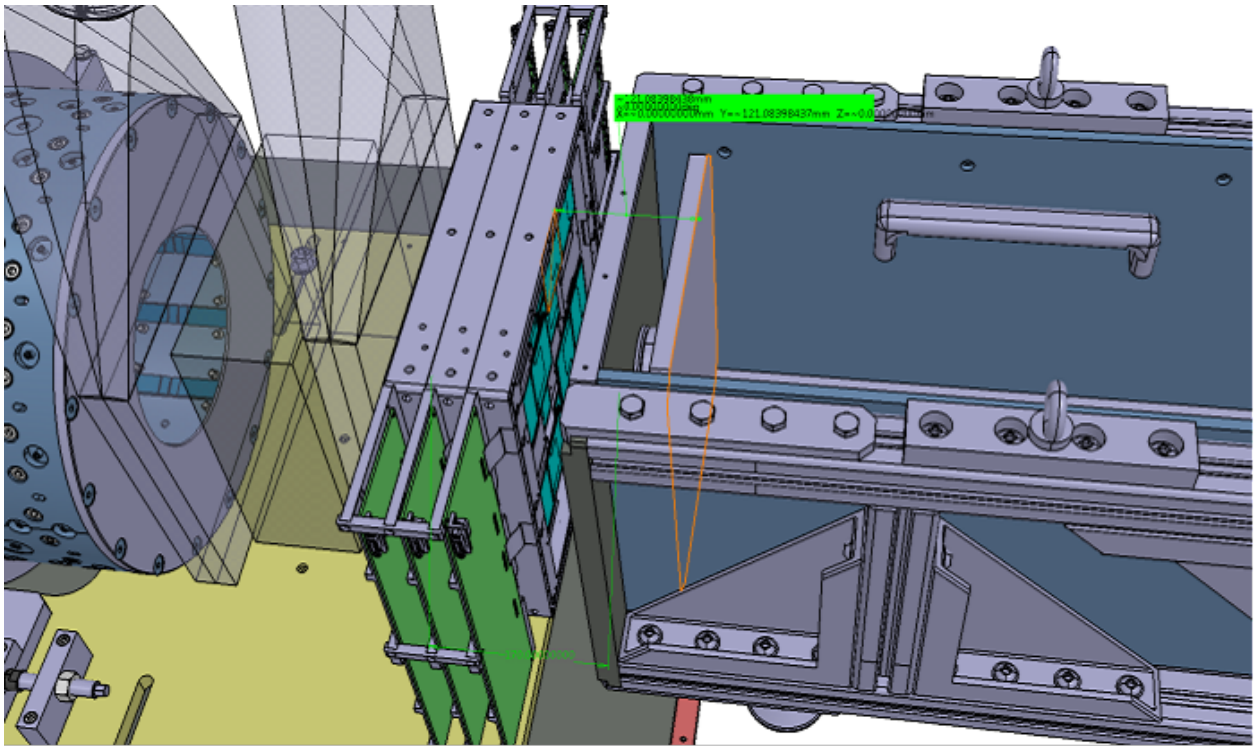

Figure E.1: CAD model of the FASERν emulsion detector and the IFT of the FASER spectrometer, with distance relevant for track matching shown in green.

Note that the extrapolated position is different for different SCT modules in the first layer of the IFT. The track matching will use the ACTS reconstructed tracks in the FASER

spectrometer; however, because emulsion reconstruction has not yet been implemented in the `calypso` offline software framework, the muon tracks in the FASERν emulsion detector that will be used for track matching are the truth tracks. The procedure for track matching is as follows;

1. Extrapolate the emulsion truth muon track to the first layer of the IFT.

2. If any ACTS reconstructed FASER tracks match within the Δr and $\Delta\theta$ thresholds, include them as forward matches.

3. Extrapolate all ACTS tracks starting in the IFT back into the FASERν emulsion detector.

4. If any truth muon tracks in the emulsion detector match within the same thresholds, include them as backward matches.

5. FASER and FASERν tracks that are both backward matches and forward matches are identified as belonging to the same particle.

6. The number of tracks with exactly one match is used to calculate the efficiency.

Observe that there is already an efficiency of $\approx 20\%$ in reconstructing the ACTS tracks in FASER due to the combinatorial Kalman filter algorithm. Track matching is relevant for the neutrino-antineutrino cross-section sensitivity study. The neutrino cross section is obtained experimentally from the following equation:

$$\sigma \equiv \frac{N}{\mathcal{L}A\epsilon_{\text{vertexID}}\epsilon_{\text{muonID}}\epsilon_{\text{emulsion}}\epsilon_{\text{tracker}}\epsilon_{\text{matching}}}. \tag{E.1}$$

Here, the quantities are defined as follows:

- N denotes the number of observed neutrino CC events.

- $\mathcal{L}$ denotes the integrated luminosity: $\mathcal{L} = 150\,\text{fb}^{-1}$ for Run 3.

- A denotes the geometrical acceptance of the detector and is different for ν_μ and $\bar{\nu}_\mu$: $A(\nu_\mu) = 0.41 \pm 0.02$, $A(\bar{\nu}_\mu) = 0.28 \pm 0.01$.

- $\epsilon_{\text{vertexID}}$ denotes the efficiency for identifying a neutrino CC vertex in the emulsion detector. This is energy dependent and is shown in Figure E.2 [2].

- ϵ_{muonID} denotes the efficiency for identifying an outgoing muon from a neutrino CC vertex, $\epsilon_{\text{muonID}} = 0.86$ [2].

- $\epsilon_{\text{emulsion}}$ denotes the emulsion tracking efficiency, $\epsilon_{\text{emulsion}} = 0.951 \pm 0.021$.

- $\epsilon_{\text{tracker}}$ denotes the FASER tracking efficiency, $\epsilon_{\text{tracker}} = 0.929 \pm 0.010$.

- $\epsilon_{\text{matching}}$ denotes the track matching efficiency between FASER and FASERν and is different for ν_μ and $\bar{\nu}_\mu$: $\epsilon_{\text{matching}}(\nu_\mu) = 0.967 \pm 0.009$, $\epsilon_{\text{matching}}(\bar{\nu}_\mu) = 0.971 \pm 0.014$.

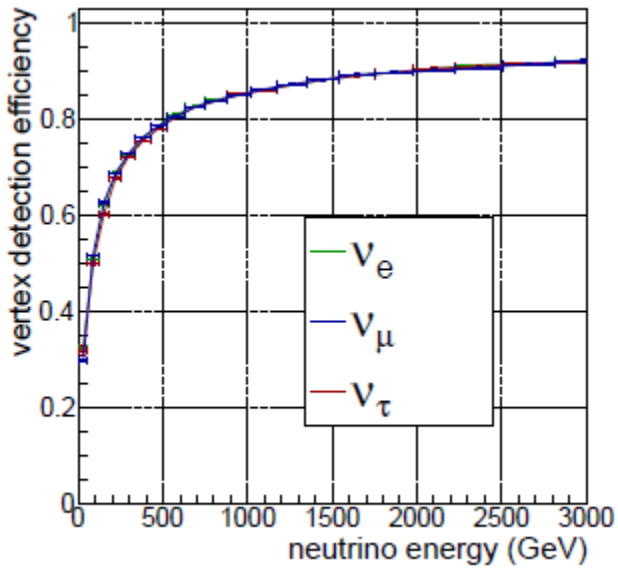

Figure E.2: $\epsilon_{\text{vertexID}}$ as a function of incident neutrino energy.

The numbers $N(\nu_\mu)_{\text{obs}}, N(\bar{\nu}_\mu)_{\text{obs}}$ of observed neutrino events are related to the numbers $N(\nu_\mu)_{\text{true}}, N(\bar{\nu}_\mu)_{\text{true}}$ of actual neutrino events (expected numbers in Table 3.1) by the charge identification rate. If $\epsilon_{\text{chargeID}}$ denotes the fraction of muons with correctly identified charges, then these are related by

$$N(\nu_\mu)_{\text{obs}} = \epsilon_{\text{chargeID}} N(\nu_\mu)_{\text{true}} + (1 - \epsilon_{\text{chargeID}}) N(\bar{\nu}_\mu)_{\text{true}}, \tag{E.2}$$

$$N(\bar{\nu}_\mu)_{\text{obs}} = (1 - \epsilon_{\text{chargeID}}) N(\nu_\mu)_{\text{true}} + \epsilon_{\text{chargeID}} N(\bar{\nu}_\mu)_{\text{true}}. \tag{E.3}$$

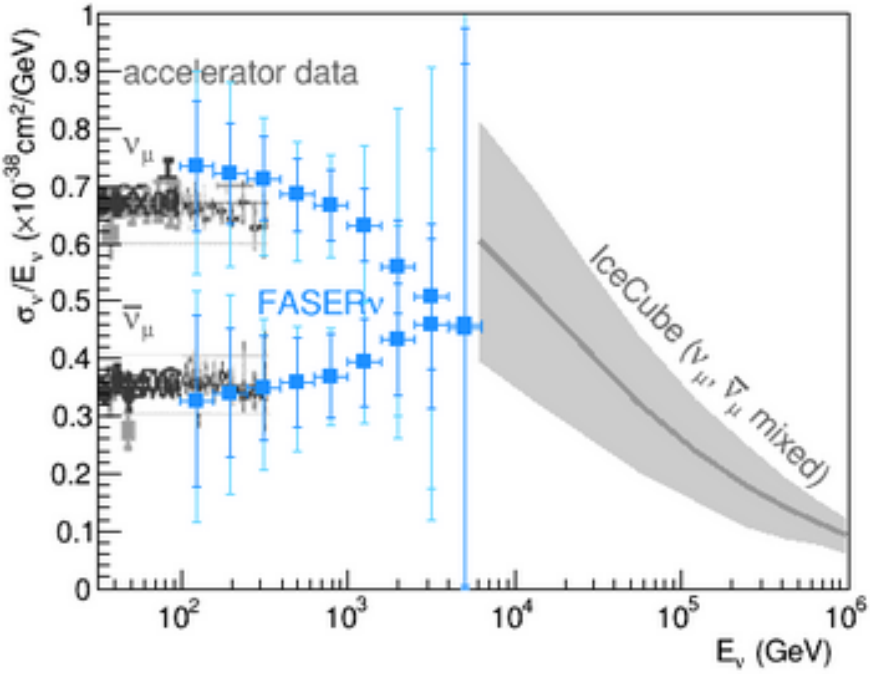

Figure E.3: Muon neutrino-antineutrino CC cross-section sensitivity as a function of energy. Error bars show statistical (dark blue) and statistical+systematic (light blue) uncertainties.

Using the charge misidentification as a source of systematic uncertainty, it is possible to get a final estimate on $\nu_\mu/\bar{\nu}_\mu$ CC cross-section sensitivity verses energy, shown in Figure E.3. Here the points plotted are derived from the expectations of ν_μ and $\bar{\nu}_\mu$ events with charge misidentification taken into account, which also gives the statistical uncertainty. The

systematic uncertainty arises from variations in the event generators as well as the charge misidentification.