UCLouvain

FACULTE DES SCIENCES

Faculté des sciences

Investigating the theoretical limit of muon track reconstruction in IceCube

Mémoire présenté en vue de l'obtention du grade académique de

Master [120] en sciences physiques

Author : Matthieu Soquay

Supervisors : Gwenhaël de Wasseige & Matti Jansson

Readers : Andrea Giammanco & Agni Bethani

School of Physics, CP3

Academic year 2025–2026

Acknowledgements

After spending an entire year working on this thesis, I would first like to sincerely thank my supervisor, Gwenhaël de Wasseige, for all her support throughout the year. Her guidance, advice, and the opportunity to work on this topic were essential in allowing me to carry out and complete this thesis.

I would also like to thank my other supervisor, Matti Jansson. I had the opportunity to work with him and to be in contact with members of the IceCube collaboration. His expertise and many valuable suggestions allowed me to approach the work from a different perspective and greatly helped me throughout the project.

This thesis is also the result of all the years I have spent studying and marks the end of an important part of my academic journey. These years would not have been the same without the people I met and the experiences I shared along the way. I would also like to offer special thanks to Guillaume, Charly, Hugues, Augustin, Edouard, Pierre, Benjamin and finally Hugo, who made these years particularly memorable.

Finally, I would like to thank the people who have supported me the most throughout all these years: my parents. Their support, from my very first year of university until today, made this work possible. I would not have been able to reach this point without them.

Contents

Acknowledgements **i**

List of Figures **vi**

List of Tables **ix**

1 Introduction **1**

1.1 Scientific motivation 1

1.2 Reconstruction as an inverse problem 2

1.3 Research question, aim and scope 3

1.4 Methodological contribution 4

1.5 Thesis outline 4

2 Neutrino physics and IceCube event topologies **5**

2.1 From the neutrino hypothesis to its observation 5

2.2 Neutrinos in particle physics 6

2.3 High-energy astrophysical neutrinos 7

2.4 Neutrino interactions and event topologies 7

3 The IceCube detector and the optical medium **9**

3.1 The IceCube Neutrino Observatory 9

3.1.1 Natural media and the South Pole ice 9

3.1.2 From AMANDA to IceCube 10

3.1.3 Scientific reach and selected results 10

3.1.4 In-ice geometry 11

3.1.5 DeepCore 11

3.2 Cherenkov-light production 14

3.3 Digital Optical Modules and pulse formation 15
3.3.1 Photomultiplier response 15
3.3.2 From a waveform to pulses 15
3.4 Photon propagation in South Pole ice 16
3.4.1 Absorption and scattering 16
3.4.2 Depth structure and dust layers 17

4 Muon track reconstruction and the TriplePandel model 19
4.1 Synthetic reconstruction problem 19
4.2 Track parametrisation and timing geometry 20
4.3 From LineFit to likelihood reconstruction 21
4.3.1 LineFit as a first estimate 21
4.3.2 Maximum-likelihood reconstruction 22
4.3.3 SPE and MPE likelihoods 22
4.4 Photon arrival-time models 23
4.4.1 Pandel parametrisation 24
4.4.2 Photon tables and spline reconstructions 24
4.5 TriplePandelMPE 25
4.6 Infinite-muon and uniform-light hypothesis 26
4.7 Evaluation principle and scope 28

5 Methodology 29
5.1 Principle and scope of the study 29
5.1.1 From a known track to a synthetic event 30
5.1.2 Reference light yield and statistical realisations 32
5.1.3 Meaning of the theoretical limit 33
5.2 Observables used to evaluate the reconstruction 33
5.3 From fixed tracks to controlled ensembles 37
5.3.1 Azimuth, corridor geometry and transverse position 37
5.3.2 Depth and the vertical DOM spacing 44
5.3.3 From selected detector sides to a radial coordinate 45
5.4 Production scans 48
5.4.1 Azimuth 49

5.4.2 Depth . . . 49
5.4.3 Zenith . . . 50
5.4.4 Detector-centred radial dependence . . . 51
5.4.5 DOM-centred local placement diagnostics . . . 52
5.5 Preservation of event-level results . . . 59

6 Results **60**
6.1 Reading conventions . . . 60
6.2 Depth dependence . . . 61
6.3 Azimuth dependence . . . 62
6.4 Zenith dependence . . . 66
6.5 Detector-centred radial dependence . . . 68
6.6 Local DOM-centred geometry . . . 72
6.6.1 Distance from the selected DOM . . . 72
6.6.2 Above–below placement . . . 74
6.7 Cross-scan synthesis . . . 76

7 Discussion **77**
7.1 Meaning of the theoretical reconstruction limit . . . 77
7.2 Global averages and local detector dependence . . . 78
7.3 Depth dependence and available optical information . . . 79
7.4 Azimuthal structure and detector geometry . . . 81
7.5 Zenith asymmetry . . . 84
7.5.1 Coordinate singularity close to the poles . . . 88
7.6 Detector-centred radial dependence and event selection . . . 90
7.7 Local DOM populations and the origin of the large tails . . . 91
7.8 Directional bias, tails and statistical interpretation . . . 96
7.9 Charge, hit multiplicity and reconstruction precision . . . 97
7.10 What is established by the present study . . . 98
7.11 Scope and limitations . . . 99

8 Conclusion **100**

A Reconstruction Metrics **103**

A.1 Directional performance 103
A.2 Spatial diagnostics 104
A.3 Detected information and event survival 105
A.4 Distribution summaries and uncertainty 105

B Complete overview dashboards **107**
B.1 Global reconstruction dashboards 107
B.2 Directional-bias dashboards 114

C Computational Workflow and Execution Details **118**
C.1 Common event-level workflow 118
C.2 Indispensable reconstruction blocks 118
C.3 How the scans reuse the workflow 120
C.4 Execution constraints and retained records 120

D AI Usage **122**

Bibliography **123**

List of Figures

1.1 Multi-messengers from hadronic accelerators. 2

2.1 Particles of the Standard Model. 6
2.2 Track-like and cascade-like neutrino events 8

3.1 Geometry of the IceCube in-ice array 12
3.2 DeepCore geometry and the dust layer 13
3.3 Indirect neutrino detection in IceCube 14
3.4 Photomultiplier and assembly of an IceCube DOM 15
3.5 From detected photons to reconstructed pulses 16
3.6 Depth-dependent optical coefficients of South Pole ice 18

4.1 Photon multiplicity and the MPE first-arrival density 23
4.2 TriplePandel reconstruction pipeline 26
4.3 Infinite-muon track hypothesis 27

5.1 Parallel tracks with finite spatial separation 34
5.2 First fixed-track azimuth scan 38
5.3 Fixed diagnostic track between string rows 39
5.4 Geometry of the parallel-offset corridor diagnostic 40
5.5 Reconstruction and charge in the parallel-offset diagnostic 41
5.6 Geometry of the broadened azimuth scan 42
5.7 Azimuth curves by transverse-offset region 43
5.8 Event-level angular-error distribution 43
5.9 Angular-error percentiles versus transverse offset 44
5.10 Sampling within a depth layer 45
5.11 Geometry of the first six-side boundary scan 46

5.12 Angular error in the exploratory six-side boundary scan 47
5.13 Total charge in the exploratory six-side boundary scan 47
5.14 Representative DOM locations used in the local diagnostics 53
5.15 DOM-centred radial and vertical sampling 55
5.16 Exploratory radial comparison at three DOM locations 57
5.17 Exploratory vertical-offset comparison at three DOM locations 57
5.18 Above–below placement diagnostic for DOM 2696 58

6.1 Angular-error percentiles as a function of the sampled depth. 61
6.2 Pooled angular-error percentiles as a function of the true azimuth. . . 62
6.3 Azimuth dependence by depth region 65
6.4 Directional and spatial reconstruction performance in the zenith scan. 67
6.5 Detector-centred radial-scan ranges 70
6.6 Event fractions and counts in the radial scan 71
6.7 DOM-centred radial result for the primary 24-DOM production. . . . 73
6.8 Vertical DOM-centred result for the same primary 24-DOM production. 75

7.1 Depth-scan light-information diagnostics 80
7.2 Top view of the string geometry and tested azimuth axes 82
7.3 Azimuthal geometry diagnostics and reconstruction response 83
7.4 Zenith-scan light-information diagnostics 85
7.5 Zenith reconstruction at three vertical sampling centres 87
7.6 Zenith p_{95} at three sampling centres 88
7.7 Expected optical information versus DOM depth in the paired zenith control 89
7.8 Local diagnostic DOM populations 92
7.9 Radial angular-error population comparison 93
7.10 Vertical angular-error population comparison 93
7.11 Per-DOM local angular-error tails by detector region 95
7.12 Radial coordinate-residual population comparison 96
7.13 Vertical coordinate-residual population comparison 96

B.1 Depth-scan overview 108
B.2 Azimuth-scan overview 109

B.3 Zenith-scan overview 110
B.4 Detector-centred radial-scan overview 111
B.5 Local radial-scan overview 112
B.6 Local vertical-scan overview 113
B.7 Azimuth-scan aligned directional bias 114
B.8 Zenith-scan aligned directional bias 115
B.9 Depth-scan coordinate residuals 116
B.10 Detector-centred radial aligned directional bias 116
B.11 Local radial-scan coordinate residuals 117
B.12 Local vertical-scan coordinate residuals 117

C.1 Common computational sequence for one production sample. 118
C.2 JAX compilation and input shapes 121

List of Tables

5.1 Final detector-wide runs, requested event counts and controlled sampling. 49
5.2 Requested event counts for the local DOM productions 56

6.1 Angular-error extrema in the final depth scan. 61
6.2 Principal numerical extrema of the pooled azimuth production. 63
6.3 Depth-resolved azimuth percentile ranges 63
6.4 Reference bins of the zenith scan 66
6.5 Reference bins of the detector-centred radial scan 68
6.6 Area-weighted isotropic references from the radial scan 69
6.7 Primary local-scan summary 76

7.1 Pooled-azimuth string-distance diagnostic 82
7.2 Paired-direction light-response control 84
7.3 Primary and extended local populations 94
7.4 Local metrics by detector region 94

A.1 Metrics retained for interpreting the reconstruction results. 106

Chapter 1

Introduction

1.1 Scientific motivation

The observation of a high-energy astrophysical neutrino population by IceCube in 2013 established neutrinos as messengers of the non-thermal Universe [1]. Once this extraterrestrial component had been identified, a central question became where these neutrinos are produced. Answering it requires reconstructed directions precise enough to associate individual events, or an excess of events, with possible astrophysical sources. Angular reconstruction is essential here, since its precision affects the sensitivity of point-source searches and the reliability of a possible source association.

Neutrinos are produced in some of the most extreme events in the Universe, and observing them with high precision can inform both neutrino physics and the astrophysical environments in which they are produced.

To study and understand these events, we use multi-messenger astronomy. Instead of relying on only one carrier of information, different messengers are combined in order to obtain a more complete and complementary picture: one messenger can provide information where another cannot. Gravitational waves, electromagnetic radiation, cosmic rays and neutrinos give different views of the same high-energy Universe [2]. In this work, I focus on neutrinos.

Neutrinos have two properties that are particularly interesting for directional studies, where the objective is to determine *where* the detected particles come from. First, neutrinos are electrically neutral. They are not deflected by electric or magnetic fields during their propagation. This is an essential advantage over charged cosmic rays, whose arrival direction does not generally point back to their production site [2].

However, the path between the production of a particle and its detection is not empty or perfectly transparent. Matter, radiation fields and dense astrophysical environments can absorb or reprocess other messengers. This is where the second important property of neutrinos becomes essential: neutrinos interact only through the weak interaction and thus interact only very rarely with matter [3]. They can consequently escape dense environments and travel over very large distances with

little attenuation.

This complementarity and the different propagation of the messengers are illustrated in Figure 1.1.

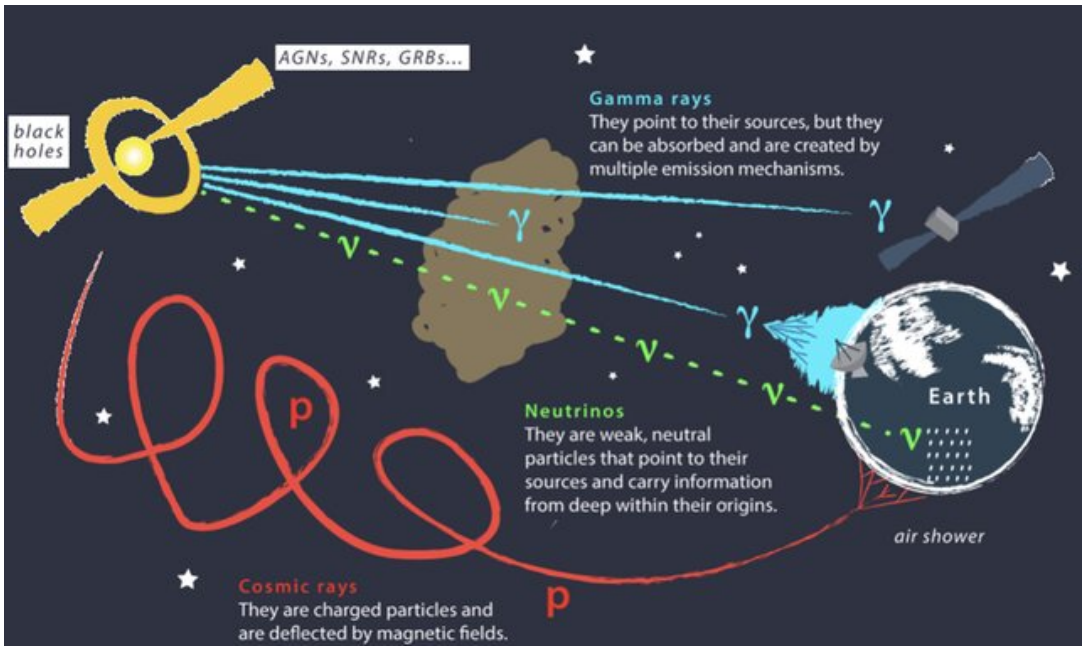


Figure 1.1: Multi-messengers from hadronic accelerators.

Source: reproduced from Montaruli [4]; original artwork credited there to J. A. Aguilar and J. Yang.

These two properties make neutrinos particularly well suited to the directional study of astrophysical sources. The same weak interaction that makes them useful messengers also makes them extremely difficult to detect. This requires a very large detection volume to have a reasonable probability of observing a neutrino interaction. This is the basic idea behind IceCube, a cubic-kilometre neutrino detector installed in the Antarctic ice. The ice provides both the required detection volume and a naturally transparent medium in which the light produced by secondary charged particles can be observed.

1.2 Reconstruction as an inverse problem

As I will explain throughout this thesis, determining where these particles come from is a reconstruction problem, and more specifically an inverse problem. IceCube does not directly observe the neutrino trajectory. In the event class considered here, a muon neutrino undergoes a charged-current interaction and produces a relativistic muon. The muon then emits Cherenkov light while travelling through the ice, and the positions, times and charges recorded by the Digital Optical Modules are used to reconstruct its track.

The muon is charged, unlike the neutrino, but its extended trajectory through the detector provides strong directional information. This is why the angular er-

ror between the true and reconstructed muon directions is the central performance metric in this work. A biased or imprecise reconstructed direction weakens the link between an observed event and a possible astrophysical source.

In this study, I reconstruct the *muon* direction from synthetic detector pulses. I do not reconstruct the neutrino–muon production angle, and the angular resolutions reported in this thesis are not neutrino angular resolutions. The true direction used throughout the analysis is the direction assigned to the simulated muon.

A reference angular value can be defined once the event population and its weighting are stated. Such a value does not, however, describe the reconstruction completely. IceCube has a discrete geometry, depth-dependent optical properties, a denser DeepCore region and a finite instrumented volume. The performance can depend on the direction and position of a track, while a good median can still hide a coherent directional bias, a long tail of poorly reconstructed events or a loss of events. These effects are precisely what I investigate in this work.

1.3 Research question, aim and scope

SplineMPE is the established infinite-muon likelihood reconstruction used in many high-energy IceCube analyses, including point-source searches. TriplePandelMPE is being developed to replace this reconstruction role with a compact, differentiable neural-network parametrisation of the photon timing response [5, 6]. Its performance matters beyond one reconstruction benchmark: if later comparisons establish both sufficient angular accuracy and a computational gain, it could affect the reconstruction used in searches for astrophysical sources. My objective is not to improve TriplePandelMPE during the thesis or to measure its execution speed, but to test it under different controlled conditions in order to determine how accurately and robustly it can reconstruct a muon track and to identify the geometrical, optical or model-dependent structures that remain in the reconstruction error.

This leads to the central question of the thesis:

> *Under matched synthetic conditions and the infinite-muon approximation, how accurately and robustly can the TriplePandel model reconstruct a muon track across the tested IceCube geometries and directional ranges, and which geometrical, optical or model assumptions structure the remaining errors?*

All events used in this study are synthetic. Event generation and reconstruction use the same detector geometry, ice description, expected light-response model, photon-detection calibration, timing convention and light-yield scale. I study a matched-model closure problem: the forward model used to generate the observations is also the model used to reconstruct them. The objective is to investigate the reconstruction limit of the continuous-emission, infinite-muon approximation inside this controlled framework rather than to reproduce the full experimental IceCube resolution. Real-data backgrounds, calibration uncertainties, an intentional mismatch between generation and reconstruction, and stochastic muon energy losses

are not part of the baseline model unless they are introduced explicitly in a dedicated test.

1.4 Methodological contribution

The main part of my work is not simply to run a large number of simulations, but to determine how the model should be tested. The first fixed scans showed that local detector geometry can strongly affect the reconstruction, for example when a track passes close to particular DOMs or follows preferred paths between strings. I progressively adapted the methodology in order to isolate one effect at a time while averaging over geometrical variables that were not being studied.

This led to the final scans in azimuth, zenith, depth, local DOM position and distance from the detector centre, together with additional light-response and geometry controls. I use angular-error percentiles, the aligned three-dimensional directional bias, detected charge, hit-DOM multiplicity and reconstruction-success fractions. Together, these observables distinguish the typical reconstruction level from a coherent displacement, a minority tail and a loss of events. The detector-centred radial production is also integrated with a declared geometrical weighting to provide one reference angular value. The detailed development of this methodology, including the preliminary tests that motivated these choices, is presented in Chapter 5.

1.5 Thesis outline

Chapter 2 introduces neutrino physics, high-energy astrophysical production, interactions and the event topologies relevant for IceCube. Chapter 3 then presents the detector, its geometry, DOMs, DeepCore, the optical properties of the ice and the detection of Cherenkov light.

Chapter 4 develops the main steps of muon-track reconstruction, from simple approaches to the likelihood methods leading to TriplePandel. Chapter 5 describes the synthetic events, the scans and the preliminary observations that shaped the final analysis.

Chapter 6 reports the final results, before Chapter 7 examines their optical, geometrical and model-related origins through targeted controls. Chapter 8 then answers the research question and discusses the main limitations and next steps. Implementation and execution details are kept in the appendices.

Chapter 2

Neutrino physics and IceCube event topologies

2.1 From the neutrino hypothesis to its observation

Pauli proposed the neutrino in 1930 to explain the continuous energy spectrum observed in beta decay without abandoning energy and angular-momentum conservation. At that stage, it was only a hypothetical light and neutral particle. Fermi later included it in his theory of beta decay and provided the first quantitative description of the process [7].

The first direct detection came in 1956, when Cowan and Reines observed reactor antineutrinos [8]. The experimental picture then became progressively richer. In 1962, an accelerator experiment showed that the neutrino associated with the muon is distinct from the electron neutrino [9]. The direct observation of the tau neutrino in 2000 completed the experimental three-flavour picture [10].

The study of naturally produced neutrinos provided a major turning point. Early solar measurements detected only about one third of the electron-neutrino flux expected from solar models [11]. Later solar and atmospheric experiments showed that the missing neutrinos had not disappeared: they had changed flavour during propagation. In particular, the Sudbury Neutrino Observatory measured a total solar-neutrino flux consistent with the solar prediction while observing that only part of it remained in the electron flavour [12]. Neutrino oscillations require non-zero mass differences, demonstrating that the minimal massless-neutrino formulation of the Standard Model is incomplete [7].

Neutrinos also became astronomical observables. The burst detected from Supernova 1987A was the first neutrino signal observed from beyond the Solar System [13]. It supported the basic picture in which a large fraction of the energy of a core-collapse supernova is carried away by neutrinos. It also showed that these particles can reveal information about an event that cannot be obtained from light alone [14].

Initially introduced as purely theoretical particles, neutrinos have become both probes of fundamental physics and messengers for studying the Universe.

2.2 Neutrinos in particle physics

In the Standard Model, as illustrated in Figure 2.1, the electron, muon and tau neutrinos belong to the lepton sector and are associated with the charged leptons e, μ and τ. They carry no electric charge and do not participate in the strong interaction. Their observable interactions are weak processes mediated by the charged bosons $W^{\pm}$ or the neutral boson Z^0. The small neutrino interaction cross section explains both the difficulty of detecting neutrinos and their ability to cross large amounts of matter [3].

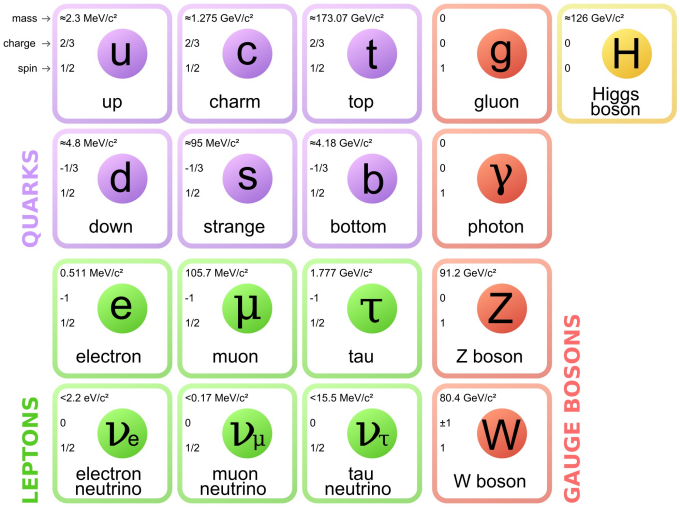


Figure 2.1: Particles of the Standard Model.

Source: reproduced from Cush via Wikimedia Commons [15], licensed under CC BY 3.0.

The minimal Standard Model treats neutrinos as massless. This description is incomplete because the flavour states produced and detected in weak interactions are not identical to the states of definite mass. Their relation through the PMNS matrix is

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U^*_{\alpha i} |\nu_i\rangle, \qquad \alpha = e, \mu, \tau, \tag{2.1}$$

where ν_i are the mass eigenstates. Equation 2.1 shows that a flavour state is a superposition of these states. As the components propagate, they acquire different phases, so the probability of observing a given flavour changes with distance and energy. The observation of neutrino oscillations requires non-zero mass differences and extends the minimal Standard Model description [7].

The oscillation parameters are not used in the reconstruction developed in this thesis. They are included here to complete the physical picture: astrophysical neutrinos arrive at Earth as a mixture of flavours, whereas the event class studied below is selected through the production of a muon in or around the detector.

At high energies, astrophysical neutrinos also extend the range available for particle-physics tests [16]. Cosmic neutrinos reach Earth with incident energies far above those achievable by accelerator-produced neutrino beams, making it possible

to study neutrino interactions in an otherwise inaccessible regime. Agreement with the Standard Model constrains possible deviations, whereas an inconsistency could indicate missing physics [17]. The same observations can also probe energetic astrophysical processes and fundamental interactions.

2.3 High-energy astrophysical neutrinos

The general multi-messenger motivation for using neutrinos was introduced in Chapter 1. Here, the focus is on the production of high-energy astrophysical neutrinos.

High-energy neutrinos are expected when accelerated protons or nuclei interact with matter or radiation near an astrophysical source. The two standard channels are hadronic pp interactions and photo-hadronic $p\gamma$ interactions. Both produce secondary mesons, especially pions and kaons, whose decays include neutrinos. A representative charged-pion chain is

$$\pi^+ \to \mu^+ + \nu_\mu, \tag{2.2}$$
$$\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu. \tag{2.3}$$

Analogous chains occur for negatively charged mesons. In this context, neutrino observations probe hadronic interactions that cannot be identified from electromagnetic radiation alone [2].

IceCube's observation of a diffuse astrophysical neutrino flux provided the first evidence for a population of high-energy neutrinos of extraterrestrial origin. Subsequent analyses measured the spectrum with several event selections, including starting-track samples that retain the directional information of muon events [2, 18]. This context motivates the reconstruction problem introduced in Chapter 1: an accurately reconstructed muon direction helps to localise the source of the parent neutrino.

2.4 Neutrino interactions and event topologies

At the energies relevant for IceCube, neutrino–nucleon interactions are separated into charged-current (CC) and neutral-current (NC) processes. In a CC interaction, the neutrino produces its associated charged lepton,

$$\begin{aligned} \nu_\ell + N &\to \ell^- + X, \\ \bar{\nu}_\ell + N &\to \ell^+ + X, \end{aligned} \tag{2.4}$$

whereas an NC interaction leaves a neutrino in the final state,

$$\begin{aligned} \nu_\ell + N &\to \nu_\ell + X, \\ \bar{\nu}_\ell + N &\to \bar{\nu}_\ell + X. \end{aligned} \tag{2.5}$$

Here, $\ell = e, \mu, \tau$, N denotes the target nucleon or nucleus and X the hadronic final state. The CC and NC processes are represented in Equations 2.4 and 2.5, respectively [3].

The particles emerging from the interaction determine the visible topology. NC events of all flavours and electron-neutrino CC events produce light patterns dominated by relatively compact hadronic or electromagnetic cascades. Tau-neutrino CC events can produce several topologies depending on their energy and decay channel, but they are not used in this work. A muon-neutrino CC interaction produces a hadronic cascade at the interaction vertex and a relativistic muon that can travel hundreds of metres or more through the ice. Cherenkov light emitted along the muon path creates an extended pattern of hit sensors, referred to as a track, as illustrated in Figure 2.2 [5, 19].

Tracks usually provide stronger directional information than cascades because the detector samples the light emission over a longer lever arm. This is the event class considered throughout the thesis. The reconstructed quantity is the muon direction, not the neutrino direction itself. The true direction used for evaluation is the direction assigned to the simulated muon. Comparing it with the reconstructed direction isolates the track-reconstruction problem and excludes the kinematic angle between the incoming neutrino and the outgoing muon.

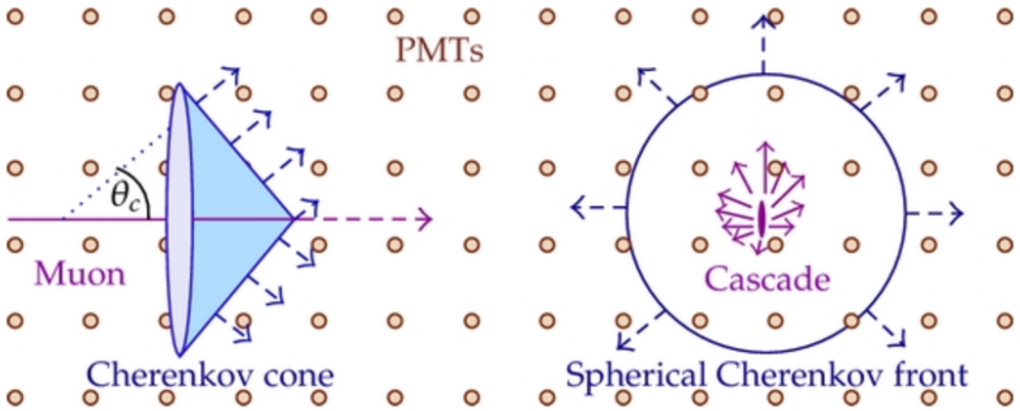


Figure 2.2: Schematic comparison of a track-like charged-current ν_μ interaction and a cascade-like neutrino interaction.

Source: adapted from Ahrens et al. [20].

The sequence described in this chapter explains why observing neutrinos has become essential. Their flavour changes revealed physics absent from the minimal Standard Model, while their propagation over astronomical distances makes them direct messengers of energetic and otherwise difficult-to-probe environments. The same weak interaction that preserves this information also makes a neutrino unlikely to interact in a detector. A very large and well-understood detection volume is required. The next chapter introduces IceCube, the observatory that makes this measurement possible, and the detector signals from which a muon track must be reconstructed.

Chapter 3

The IceCube detector and the optical medium

3.1 The IceCube Neutrino Observatory

Chapter 2 showed why neutrinos are valuable messengers, but also why they are difficult to observe. Their small interaction probability requires a very large target volume, while their direction can only be inferred from the secondary charged particles produced in the rare interactions that do occur. IceCube was developed with these two constraints in mind. This chapter describes the detector and explains how it records the light produced by muons following neutrino interactions.

3.1.1 Natural media and the South Pole ice

As discussed in the previous chapter, the detection medium must provide enough target material for a neutrino interaction to occur. Building and maintaining an artificial transparent volume at the cubic-kilometre scale would not be realistic. This is why large neutrino telescopes use naturally available water or ice, which acts both as the interaction target and as the medium in which Cherenkov light propagates. Deep water generally scatters light less than glacial ice, preserving photon timing and allowing very good angular resolution, but it absorbs more light than the deep Antarctic ice. The longer absorption length of the ice allows photons to remain detectable over larger distances, partly compensating for its stronger scattering. Water also contains optical background from radioactive potassium and time-dependent bioluminescence, while detector lines can move with water currents. In glacial ice, the modules remain fixed and there are no currents or bioluminescence. Its main disadvantages are stronger scattering and the depth-dependent impurities discussed later in this chapter. Water and ice thus offer different compromises; neither medium is preferable in every respect [21].

The South Pole ice is particularly suitable because it provides a stable, transparent and extremely large natural volume. The detector is installed deep enough for the overlying ice to suppress many low-energy particles produced in cosmic-ray air showers. Penetrating atmospheric muons nevertheless remain the dominant

background: at trigger level, they outnumber atmospheric-neutrino interactions by roughly six orders of magnitude [22]. Depending on the analysis, this background can be reduced by using the Earth as a filter for up-going events, by treating the outer in-ice layers as a veto, by using information from the IceTop surface array, or by selecting sufficiently energetic and well-reconstructed events. The depth and size of IceCube are essential, but they do not remove the need for accurate event reconstruction.

3.1.2 From AMANDA to IceCube

Early projects established the feasibility of instrumenting deep water or ice, and the Antarctic Muon and Neutrino Detector Array (AMANDA) demonstrated that the South Pole ice could support a large optical detector. IceCube extended this approach to the cubic-kilometre scale. Construction was completed at the end of 2010 and commissioning followed in 2011, providing the scale required for its astrophysical-neutrino programme [23].

3.1.3 Scientific reach and selected results

The discovery of the diffuse astrophysical flux introduced in Chapter 2 was the first major step [2]. The observatory then began to connect parts of that flux to the sky. The neutrino alert associated with the flaring blazar TXS 0506+056 [24], the evidence for emission from the active galaxy NGC 1068 [25], and the observation of high-energy neutrinos from the Galactic plane [26] illustrate this progression from detecting an astrophysical flux to studying individual sources and the Milky Way itself.

Beyond its role as a neutrino observatory, IceCube also provides a unique environment for particle-physics studies. DeepCore measurements constrain atmospheric-neutrino oscillations, while the full detector is used to test neutrino interactions and to search for effects such as sterile neutrinos or dark-matter signatures [27]. In 2021, IceCube reported a particle shower consistent with the Glashow resonance [28], produced when an electron antineutrino interacts with an electron at a resonant energy of about 6.3 PeV. The process had been predicted in 1960, and its observation tested the Standard Model in a previously unexplored high-energy neutrino regime.

Obtaining these results first requires the events to be reconstructed from the observed light. For the track-like events studied in this work, the angular resolution determines how precisely an observed neutrino can be associated with a potential astrophysical source. Reconstruction models have an impact beyond a purely technical improvement: by setting the attainable directional precision, they contribute to the observational limits of IceCube for both particle-physics measurements and astrophysical-source studies. The rest of this chapter describes the detector information needed to understand that reconstruction problem.

IceCube detects neutrinos indirectly through Cherenkov emission, the same physical phenomenon responsible for the characteristic blue glow observed in nuclear reactors. For the track-like events considered here, the light is produced mainly by

secondary muons and recorded by optical modules in the in-ice array. IceCube also includes the IceTop surface array and the denser DeepCore sub-array within the detector. These components are described below, although only the in-ice detector geometry is used in the simulations presented in this thesis.

3.1.4 In-ice geometry

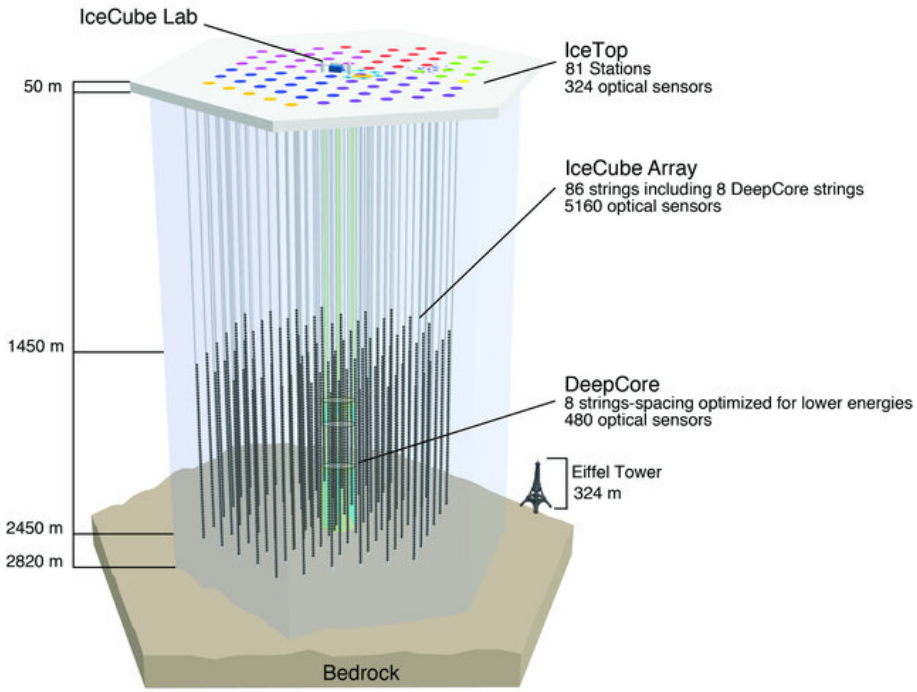
The completed in-ice array contains 5160 Digital Optical Modules (DOMs) deployed on 86 vertical strings between depths of approximately 1450 m and 2450 m. Each string carries 60 DOMs. The full geometry is shown in Figure 3.1. In the main array, neighbouring modules are separated vertically by about 17 m, and the characteristic horizontal string spacing is about 125 m, as illustrated in Figure 3.1b. Viewed from above, the strings form an approximately hexagonal footprint on a triangular grid, as seen in the same figure [23, 29].

The detector samples the light field only at discrete positions. Two tracks with the same direction and light yield can illuminate different sets of DOMs when their positions change, and the sensor coverage becomes asymmetric near the outer strings. This discrete geometry is one of the effects explored by the scans introduced in Chapter 5.

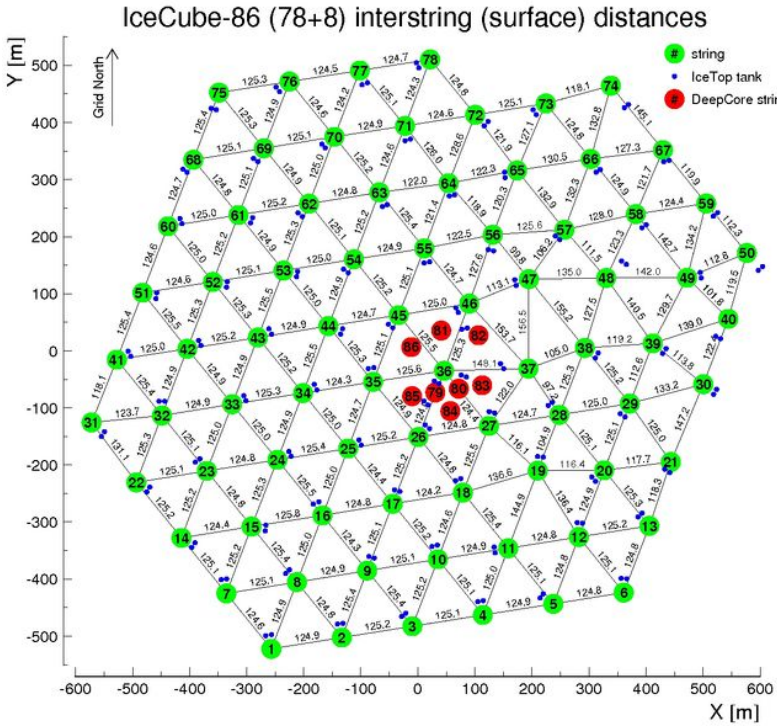
3.1.5 DeepCore

DeepCore occupies the lower central region of the array and has the highest DOM density. This denser instrumentation improves the detector response at low energies. It combines eight dedicated strings with seven nearby standard strings, as illustrated in Figure 3.2. The characteristic horizontal spacing is approximately 72 m, while the lower DOMs on the DeepCore strings are separated vertically by about 7 m. In the real detector, most modules on the dedicated DeepCore strings use high-quantum-efficiency (HQE) photomultipliers, whose photon-detection efficiency is higher than that of standard modules [22, 23, 29].

The DeepCore representation in this synthetic study, including the treatment of standard and HQE modules, is specified in Chapter 5.



(a) Three-dimensional detector layout.



(b) Top view of the string layout and inter-string distances.

Figure 3.1: Overview of the IceCube detector and top view of the in-ice string layout. The denser group of strings near the centre is DeepCore.

Sources: (a) reproduced from Collaboration et al. [23]; (b) reproduced from the IceCube Wiki [30].

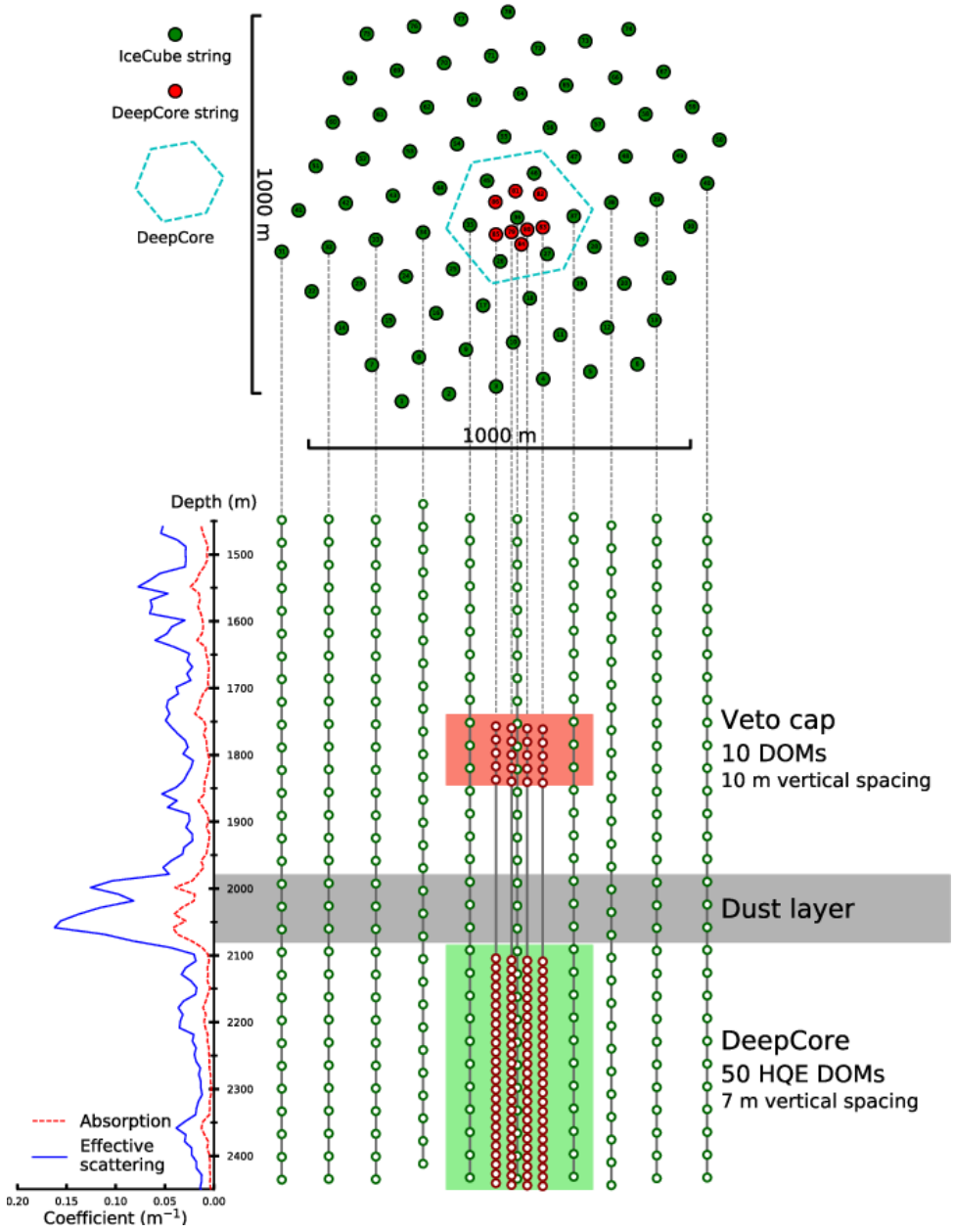


Figure 3.2: Schematic placement of the denser DeepCore modules relative to the main array and the dust-rich ice layer. The real detector distinction between standard and HQE modules shown in such detector descriptions is not included as a separate response in the simulations of this thesis.

Source: reproduced from Abbasi et al. [22].

3.2 Cherenkov-light production

When a charged particle moves through a dielectric medium, it polarises the surrounding material. In the case where its speed exceeds the phase velocity of light in that medium, the emitted disturbances add coherently and form what we call Cherenkov radiation [31]. The condition is

$$\beta n > 1, \tag{3.1}$$

where $\beta = v/c$ and n is the refractive index. The opening angle of the Cherenkov cone satisfies

$$\cos\theta_{\mathrm{C}} = \frac{1}{\beta n}. \tag{3.2}$$

For a relativistic muon, light is emitted along an extended trajectory rather than from a single point, as illustrated in Figure 3.3. The geometrical relation between the track, the emission point and a DOM determines the earliest possible arrival time of an unscattered photon [5].

The primary muon is not the only source of light as it travels through the detector. Ionisation and radiative energy losses, such as bremsstrahlung, pair production and photonuclear interactions, generate additional charged particles that also emit Cherenkov light. Their contribution becomes more important as the muon energy increases, resulting in a larger total light yield.

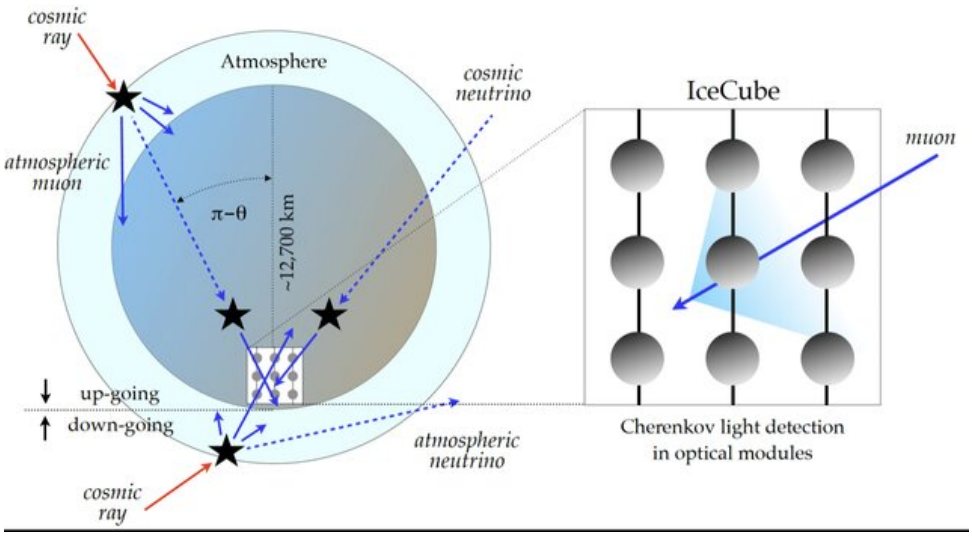


Figure 3.3: Principle of indirect neutrino detection in IceCube. Charged secondary particles emit Cherenkov light that is recorded by the optical modules.

Source: reproduced from Montaruli [4].

3.3 Digital Optical Modules and pulse formation

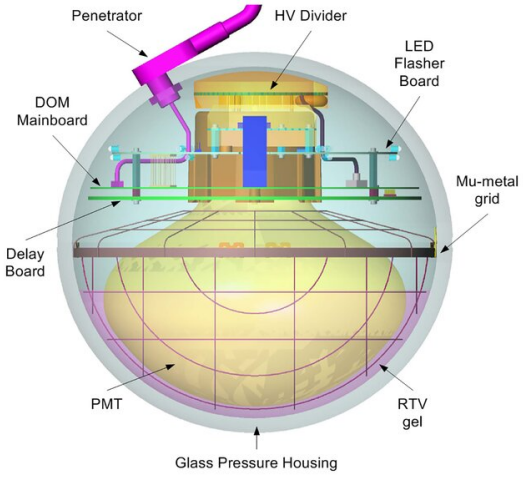
3.3.1 Photomultiplier response

The Cherenkov light is recorded by the Digital Optical Modules (DOMs) distributed throughout the instrumented volume. Each DOM is an autonomous optical sensor enclosed in a pressure-resistant glass sphere, as illustrated in Figure 3.4. It contains a downward-facing 10-inch photomultiplier tube (PMT) and the electronics required to record and transmit the signal [23].

When a photon reaches the photocathode, it can release a photoelectron. A sequence of dynodes then amplifies this electron into an electrical signal. The timing and charge properties of this response were characterised before deployment and are included in the detector calibration and simulations [32]. The signal is digitised, assigned a precise timestamp and transmitted to the surface. Each DOM thus provides both the amount of detected light and its arrival time, two quantities used directly in event reconstruction. The relative timing between DOMs is regularly calibrated across the detector [23, 33].



(a) Ten-inch photomultiplier tube of the type housed inside a DOM.



(b) DOM assembly with the downward-facing photomultiplier and electronics.

Figure 3.4: Photomultiplier and schematic assembly of an IceCube Digital Optical Module. The orientation of the photocathode gives the sensor response a vertical dependence, which motivates the above–below placement checks introduced in Chapter 5.

Sources: (a) reproduced from Hamamatsu Photonics [34]; (b) reproduced from Schatto [5].

3.3.2 From a waveform to pulses

Several photons can contribute to the same recorded waveform. The calibrated waveform is processed into a series of pulses, each characterised by an arrival time and an integrated charge, as illustrated in Figure 3.5. Photons arriving within a short

interval may contribute to the same pulse, whereas sufficiently separated signals can be resolved individually. The resulting pulse series is a compact representation of the detected light and is the form of event information used by track-reconstruction methods [5, 35].

For each DOM i, the reconstruction relies mainly on its known position $\mathbf{r}_i$, the pulse time t_i and the measured charge Q_i. The timing pattern is particularly sensitive to the track geometry: photons following direct or weakly scattered paths tend to arrive close to the geometrical expectation, whereas more strongly scattered photons arrive later [5, 35]. The charge describes the amount of detected light and enters the multiple-photoelectron likelihood introduced in Chapter 4.

In the real detector, additional conditions are applied before these signals are used. Waveform recording is initiated when the PMT signal crosses a discriminator threshold, typically around 0.25 photoelectron, while local-coincidence and higher-level trigger conditions help suppress isolated noise hits and reduce the amount of recorded data [23, 33, 35]. Noise is part of the real detector response, whereas the simulations used in this work adopt a simplified definition of a hit, as discussed in Chapter 5.

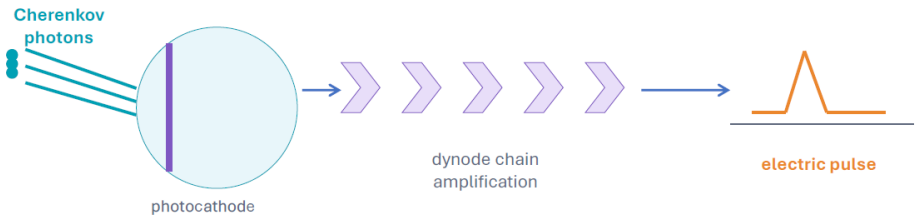


Figure 3.5: Schematic transformation of detected photons into a digitised waveform and a reconstructed pulse series.

3.4 Photon propagation in South Pole ice

3.4.1 Absorption and scattering

After their emission, Cherenkov photons propagate through the glacial ice before reaching a DOM. During this propagation, two optical effects are particularly important: absorption and scattering. Absorption removes photons from the light field and is characterised by an absorption length λ_a, while scattering changes their propagation direction and increases the distance they travel. If λ_b denotes the geometric scattering length and $g = \langle \cos \vartheta \rangle$ the mean cosine of the scattering angle, an effective scattering length can be defined as

$$\lambda_{\text{eff}} = \frac{\lambda_b}{1 - g}. \tag{3.3}$$

Both effects depend on the photon wavelength and on depth, since the concentration and composition of impurities vary throughout the ice [36].

These two effects do not influence the detected light in the same way. Absorption mainly reduces the number of photons that can reach a DOM, whereas scattering affects both their direction and their arrival time. A scattered photon generally follows a longer path and can contribute to the delayed part of the photon arrival-time distribution. As a result, the signal expected at a DOM cannot be described by a single deterministic arrival time. Track reconstruction must instead rely on a probability density for the photon arrival times, conditioned on the track–DOM geometry and on the optical properties of the ice.

3.4.2 Depth structure and dust layers

The glacial ice in which IceCube is embedded is not optically uniform. Its properties vary significantly with depth, mainly because the concentration and composition of impurities changed during the formation of the different ice layers. This produces a pronounced vertical structure in the measured optical coefficients. In particular, a dust-rich layer is found around 2000–2100 m below the surface, where both the absorption and effective scattering lengths are shorter than in the clearer ice below. DeepCore was placed largely in this deeper and more transparent region [22, 36].

This depth dependence is particularly relevant for reconstruction and will be examined in more detail in Chapter 5, where its influence on the reconstruction performance is studied.

The optical properties of the ice also depend on wavelength. Around 2400 m depth, the average effective scattering length is approximately 50 m and the average absorption length approximately 155 m for 400 nm light. At this wavelength, absorption due to dust is relatively weak, the ice is close to its maximum transparency, and the DOM sensitivity is also near its maximum. For this reason, 400 nm provides a useful reference wavelength for describing light propagation in IceCube [22]. Averaged over depths below 2100 m, the ice was estimated to be about 40–50% clearer, in terms of absorption and scattering lengths, than the ice between 1500 and 2000 m [22]. Figure 3.6 illustrates this depth dependence through the corresponding optical coefficients at 400 nm.

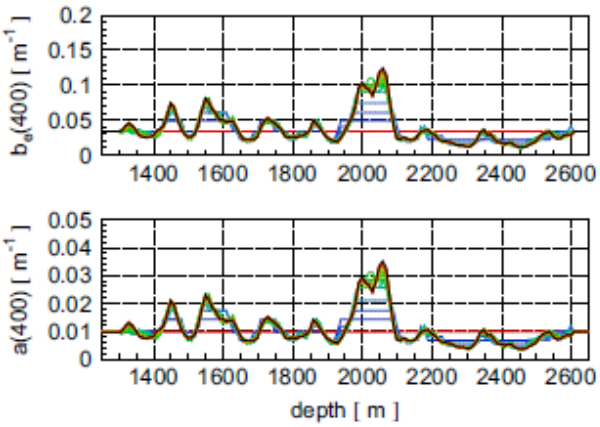


Figure 3.6: Depth dependence of the effective scattering (b_e) and absorption (a) coefficients of South Pole ice at 400 nm.

Source: adapted from Aartsen et al. [36].

This chapter has shown both why IceCube is well suited for neutrino detection and why reconstructing the observed events remains challenging. Its large instrumented volume makes rare neutrino interactions observable, but the original muon direction must then be inferred from a discrete set of pulse times and charges after the Cherenkov light has propagated through scattering and depth-dependent ice. The scientific results introduced at the beginning of this chapter ultimately rely on our ability to convert these detector signals into physical information.

The detector description developed here provides the forward picture needed for the rest of this thesis: a muon trajectory produces light that propagates through the ice and is eventually recorded by a finite set of DOMs as pulse times and charges. Chapter 4 now considers the inverse problem: how these observables can be used to reconstruct the original track, from LineFit to the TriplePandelMPE model used in this work. Chapter 5 then introduces the strategy used to test this reconstruction and the quantities required to interpret its performance.

Chapter 4

Muon track reconstruction and the TriplePandel model

The reconstruction of an event in IceCube relies on the information recorded by the DOMs, in particular their positions, the arrival times of the detected light, and the associated charges. These observables provide the information needed to infer the trajectory of the particle that produced the signal. Different reconstruction approaches have been developed for this purpose, ranging from simple and fast methods to more detailed likelihood-based models that require significantly more computation. This chapter introduces these approaches before focusing on TriplePandelMPE, which is intended to replace the present SplineMPE infinite-muon reconstruction in applications where directional precision is central, including point-source searches [6].

4.1 Synthetic reconstruction problem

It is important to mention that no real IceCube event is reconstructed in this study. Instead, each event is generated synthetically from a muon track whose parameters are known. The generating parameters are denoted by

$$\Theta_{\text{true}} = (\mathbf{r}_{0,\text{true}}, t_{0,\text{true}}, \theta_{\text{true}}, \phi_{\text{true}}) \,. \tag{4.1}$$

The principle of the reconstruction is straightforward. A forward model converts this true track into a synthetic detector response using the detector geometry and optical description introduced in Chapter 3. The resulting event consists of simulated pulses associated with the DOMs, including their times and charges. However, the reconstruction algorithm is never given Θ_{true}: recovering these parameters from the detector response is precisely the problem being tested.

Starting only from these simulated observables, the reconstruction searches for a second track,

$$\Theta_{\text{reco}} = (\mathbf{r}_{0,\text{reco}}, t_{0,\text{reco}}, \theta_{\text{reco}}, \phi_{\text{reco}}) \,, \tag{4.2}$$

that maximises the likelihood of the observed event. In these two parameter vectors, $\mathbf{r}_0$ is a reference position on the track, t_0 is the time at that position, and θ and ϕ are

the zenith and azimuth directions. The subscripts "true" and "reco" distinguish the values used to generate the event from those returned by the fit. Equivalently, the numerical optimisation minimises the negative log-likelihood. The true parameters are retained only for the later evaluation of the reconstruction. This separation between the generating track and the fitted track is central to the analysis: the reconstruction has access to the synthetic detector signal, but not to the parameters used to produce it.

This separation can be written as

$$\Theta_{\text{true}} \xrightarrow{\text{event generation}} D_{\text{synthetic}} \qquad D_{\text{synthetic}} \xrightarrow{\text{likelihood fit}} \widehat{\Theta}_{\text{reco}}. \tag{4.3}$$

Only the synthetic detector information $D_{\text{synthetic}}$ is passed to the fitter; Θ_{true} is not one of its inputs.

Even in a matched-model configuration, the inverse problem is not trivial. Photon detection and the resulting pulse pattern are stochastic, the detector samples the light field only at discrete DOM positions, and the likelihood must be optimised numerically. Different pulse realisations can lead to slightly different reconstructed tracks even when the initial muon geometry is identical.

4.2 Track parametrisation and timing geometry

Over the dimensions relevant to the detector, a high-energy muon track can be approximated as a straight line [5]. The track hypothesis used for reconstruction is described by a space–time reference point and a propagation direction,

$$\Theta = (\mathbf{r}_0, t_0, \theta, \phi). \tag{4.4}$$

Its spatial trajectory is written as

$$\mathbf{r}(s) = \mathbf{r}_0 + s\,\hat{\mathbf{u}}(\theta, \phi), \tag{4.5}$$

with

$$\hat{\mathbf{u}}(\theta, \phi) = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}. \tag{4.6}$$

The zenith angle θ is measured from the positive z-axis, while the azimuth ϕ is measured in the xy-plane from the positive x-axis towards the positive y-axis.

Here, $\mathbf{r}_0 = (x_0, y_0, z_0)$ is a chosen reference point on the track, t_0 is the corresponding reference time, and s is the signed distance along the line. For a through-going muon, this reference point does not have to correspond to a neutrino-interaction vertex, particularly in the infinite-muon approximation introduced later in this chapter. Shifting the reference point along the track changes both the parametrisation and its reference time. The physical trajectory itself remains unchanged.

For a DOM located at $\mathbf{r}_i$, the longitudinal coordinate along the track is

$$s_i = (\mathbf{r}_i - \mathbf{r}_0) \cdot \hat{\mathbf{u}}, \tag{4.7}$$

and the perpendicular track–DOM distance is

$$d_{\perp,i} = |(\mathbf{r}_i - \mathbf{r}_0) - s_i \hat{\mathbf{u}}| \,. \tag{4.8}$$

These quantities describe the position of a DOM relative to the trial track and are directly relevant to the photon-timing model.

The measured pulse time can then be compared with the geometrical arrival time expected for unscattered Cherenkov light through the time residual

$$t_{\mathrm{res},i} = t_{\mathrm{hit},i} - [t_{\mathrm{geo},i}(\Theta) + t_0] \,. \tag{4.9}$$

Here, $t_{\mathrm{geo},i}(\Theta)$ represents the geometrical time difference between the track reference point and the expected photon arrival at DOM i. It accounts for the propagation of the muon along the track to the Cherenkov emission point and the subsequent propagation of the photon through the ice, while t_0 sets the reference time of the track.

Direct or weakly scattered photons tend to populate the region close to zero residual, whereas photons following longer, more strongly scattered paths form the positive tail of the distribution [5, 35].

4.3 From LineFit to likelihood reconstruction

4.3.1 LineFit as a first estimate

LineFit provides one of the simplest ways to obtain an initial estimate of the track direction. It describes the progression of the observed hit pattern through a linear relation between DOM positions and pulse times, without explicitly modelling the Cherenkov emission geometry or the propagation of light through the ice. For DOM positions $\mathbf{r}_i$ and pulse times t_i, the corresponding least-squares problem can be written as

$$\chi^2(\mathbf{r}_0, \mathbf{v}, t_0) = \sum_{i=1}^{N_{\mathrm{hit}}} \|\mathbf{r}_i - \mathbf{r}_0 - \mathbf{v}(t_i - t_0)\|^2 \,, \tag{4.10}$$

Here, N_{hit} is the number of hit DOMs included in the fit, $\mathbf{r}_i$ and t_i are the position and pulse time of DOM i, $\mathbf{r}_0$ and t_0 are the fitted space–time reference, and $\mathbf{v}$ is the effective velocity of the observed light pattern. The direction of $\mathbf{v}$ provides the first estimate of the track direction [5].

The main advantage of LineFit is its simplicity and speed, which makes it useful as a starting point for more detailed reconstructions. However, the fitted velocity describes the progression of the observed light pattern rather than the physical velocity of the muon itself. In addition, delayed photons are not treated through any description of photon propagation in the ice. LineFit can provide a useful first estimate, but a more complete use of the timing information requires a model of the expected photon arrival times.

4.3.2 Maximum-likelihood reconstruction

This additional information can be included through a likelihood-based reconstruction. For the synthetic events considered here, the detector information can be represented as

$$D = \{(\mathbf{r}_i, t_i, Q_i)\}_{i=1}^{N_{\text{hit}}}, \tag{4.11}$$

Here, D denotes the complete synthetic event passed to the fit, N_{hit} is its number of hit DOMs, and each tuple contains the DOM position $\mathbf{r}_i$, first-pulse time t_i and sampled integer charge proxy Q_i. In the synthetic generator, Q_i is drawn from a Poisson distribution with expectation $s_E\mu_i$, where s_E is the common light-yield scale and μ_i is the unscaled expected response. A DOM contributes to the event when $Q_i > 0$. During reconstruction, this observed charge is used to determine the effective MPE multiplicity rather than through a separate charge-likelihood term.

For a trial track Θ, the likelihood evaluates how well the predicted photon arrival-time distributions reproduce the observed pulse times. The reconstructed track is obtained by maximising this likelihood, or equivalently by minimising its negative logarithm,

$$\widehat{\Theta} = \arg\max_{\Theta} \mathcal{L}(\Theta; D) = \arg\min_{\Theta} \left[-\ln \mathcal{L}(\Theta; D)\right]. \tag{4.12}$$

If the selected DOM contributions are treated as conditionally independent for a fixed track hypothesis, the event likelihood can be written schematically as

$$\mathcal{L}(\Theta \mid D) = \prod_{i=1}^{N_{\text{hit}}} p_i^{\text{MPE}}\left(t_i \mid n_i, \mathbf{r}_i, \Theta\right), \tag{4.13}$$

Here, $\mathcal{L}(\Theta \mid D)$ is the event likelihood for trial track Θ, p_i^{MPE} is the first-arrival timing density for DOM i, and n_i is obtained from Q_i using the clipping convention implemented in the reconstruction. The exact definition of this multiplicity, together with the likelihood floor and numerical objective, is given in Section C.2. The reconstruction depends directly on the photon arrival-time distribution predicted for each track–DOM geometry [5].

4.3.3 SPE and MPE likelihoods

The simplest timing likelihood is the single-photoelectron (SPE) case, where the first pulse time observed at a DOM is evaluated using a single-photon probability density $p_i(t \mid \mathbf{r}_i, \Theta)$. Its cumulative distribution is

$$P_i(t \mid \mathbf{r}_i, \Theta) = \int_{-\infty}^{t} p_i(t' \mid \mathbf{r}_i, \Theta)\, \mathrm{d}t'. \tag{4.14}$$

When several photons are detected by the same DOM, the situation changes because the reconstruction uses the arrival time of the first one. If n_i photons are independently detected, the corresponding first-arrival density becomes

$$p_i^{\mathrm{MPE}}(t_i \,|\, n_i, \mathbf{r}_i, \Theta) = n_i \, p_i(t_i \mid \mathbf{r}_i, \Theta) \, [1 - P_i(t_i \mid \mathbf{r}_i, \Theta)]^{n_i - 1} \,. \tag{4.15}$$

For $n_i = 1$, this expression naturally reduces to the SPE case. As the number of detected photons increases, the probability that at least one of them arrives early also increases, shifting the first-arrival distribution towards earlier times. This effect is illustrated in Figure 4.1. In the reconstruction used here, n_i is obtained from the sampled charge Q_i, so the charge acts as a proxy for the detected photoelectron multiplicity and directly modifies the timing likelihood [5, 35].

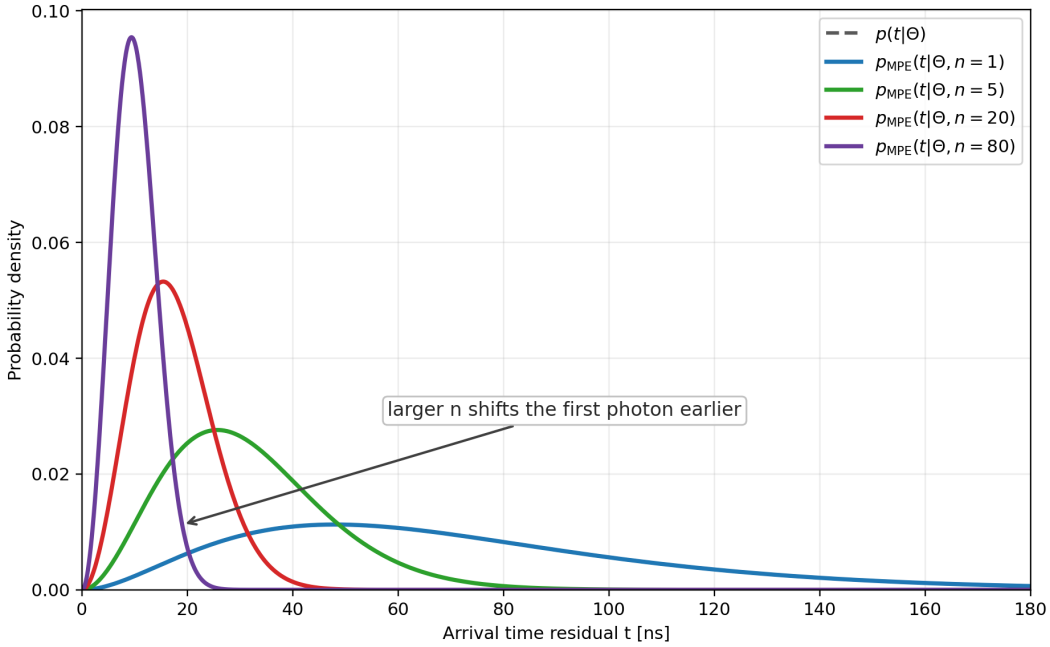


Figure 4.1: Effect of the photon multiplicity N on the MPE first-arrival density. As N increases, the first detected photon is statistically expected to arrive earlier.

4.4 Photon arrival-time models

We have now seen how the reconstruction problem leads to the need for a photon arrival-time probability density. From this point, several approaches can be used to obtain these PDFs, which form one of the central ingredients of the reconstruction. In practice, the principle remains relatively straightforward: for each trial track, the expected PDFs are evaluated and passed to the likelihood, while the fitter searches for the track that best describes the observed pulses. What mainly changes from one reconstruction to another is the model used to obtain these PDFs, together with the level of detail it can describe and the computational cost required to evaluate it. In this section, we introduce the main approaches relevant to this work.

4.4.1 Pandel parametrisation

One way to evaluate the photon arrival-time PDF is to use an analytical approach. The Pandel function was introduced by Pandel [37] as an analytical model of the time-residual distribution in a scattering medium. Its gamma-like shape depends on the track–DOM distance and on effective absorption and scattering parameters. At short distances, the density is concentrated near zero residual; as the distance increases, the distribution broadens and the scattered-light tail becomes more prominent. Practical implementations also assign a non-zero probability to small negative residuals associated with the timing resolution [5].

We will not develop the complete Pandel expression here because it is not directly evaluated in the reconstruction used in this study. However, it provides a useful example of how the photon timing distribution can be described from a relatively small number of analytical parameters.

The PandelSPE and PandelMPE models use the same underlying time distribution but differ in their treatment of photon multiplicity. These reconstructions are computationally inexpensive, but a small number of parameters cannot easily reproduce all the local variations introduced by the depth-dependent optical properties of the ice [5].

4.4.2 Photon tables and spline reconstructions

A more detailed approach is to simulate photon propagation for a multidimensional grid of source and receiver configurations. The expected photon yield and arrival-time distributions are stored in large tables and interpolated with splines during reconstruction. This retains a more detailed description of the detector geometry and ice properties without requiring a new photon-propagation simulation for every likelihood evaluation [5, 19].

SplineReco-type methods use these interpolated photon tables within an SPE or MPE likelihood. The SegmentedSplineReco model goes one step further by also introducing a more detailed description of the physical high-energy muon. Localised energy losses are represented as cascade-like sources distributed along the track, which, as discussed in Chapter 3, provides a representation closer to the actual non-uniform light production of a high-energy muon. This can improve the description of the event, but it also requires a more complex likelihood and additional fitted quantities [19].

These approaches form a progression in which more information about photon propagation or light production can be retained, generally at the cost of additional computation and implementation complexity. This is important because the timing PDF is not evaluated only once: it must be obtained repeatedly for many DOMs and for every track hypothesis tested by the fitter.

This provides the context for TriplePandel. Rather than changing the general principle of the reconstruction, it addresses the way these photon arrival-time PDFs are obtained. The production scans themselves use the uniform infinite-muon model described below.

4.5 TriplePandelMPE

TriplePandel was developed in this context to provide a faster replacement for the infinite-muon SplineMPE reconstruction while retaining the detailed dependencies contained in the underlying photon-propagation description [6]. Because SplineMPE is an important directional reconstruction in point-source analyses, preserving angular information while reducing the repeated PDF-evaluation cost has a direct scientific impact. TriplePandel uses a neural network trained to reproduce the timing distributions relevant to the IceCube geometry. In the terminology used below, *TriplePandelNET* denotes this neural-network parametrisation, while *TriplePandelMPE* denotes the complete MPE reconstruction that uses it; *TriplePandel* refers to the framework more generally.

More specifically, TriplePandel provides a compact and differentiable parametrisation of photon arrival-time distributions for the infinite-muon geometry. Instead of interpolating a large photon table at every likelihood evaluation, TriplePandelNET maps the geometrical variables describing a trial track and a DOM to the parameters of an analytical mixture distribution [6].

For DOM i, the predicted density can be written schematically as

$$p_i(t \mid \Theta) = \sum_{m=1}^{M} \alpha_{im}(\Theta) f_m\left(t; \boldsymbol{\psi}_{im}(\Theta)\right), \tag{4.16}$$

where M is the number of mixture components, the mixture weights satisfy $\alpha_{im} \geq 0$ and $\sum_m \alpha_{im} = 1$, and m labels one component. The functions f_m are Gupta, or exponentiated-exponential, components, and $\boldsymbol{\psi}_{im}$ contains their parameters. Because both the PDF and CDF are available analytically, the output can be inserted directly into the MPE first-arrival expression of Eq. (4.15) [6].

The neural network is not the reconstruction algorithm by itself. Its role is to provide a fast evaluation of the timing PDF for each candidate track–DOM configuration. Once this distribution has been obtained, the same reconstruction principle described previously applies: the likelihood compares the expected timing distributions with the observed pulses, while the optimiser varies the track parameters until the best-fitting hypothesis is found.

The track–DOM geometry is evaluated for each candidate hypothesis, while the learned parametrisation contains the ice and detector-response behaviour represented in the training sample. In this way, repeated access and interpolation of large photon tables is replaced by a direct evaluation of the PDF parameters through the network. This is particularly useful because this operation must be repeated many times throughout a single reconstruction.

The network must nevertheless be used within the detector geometry, sensor response and optical model for which it was developed. A substantial change in these conditions would require a new validation and, if necessary, an adaptation or retraining of the network.

For each trial track, the reconstruction proceeds through the following steps, as

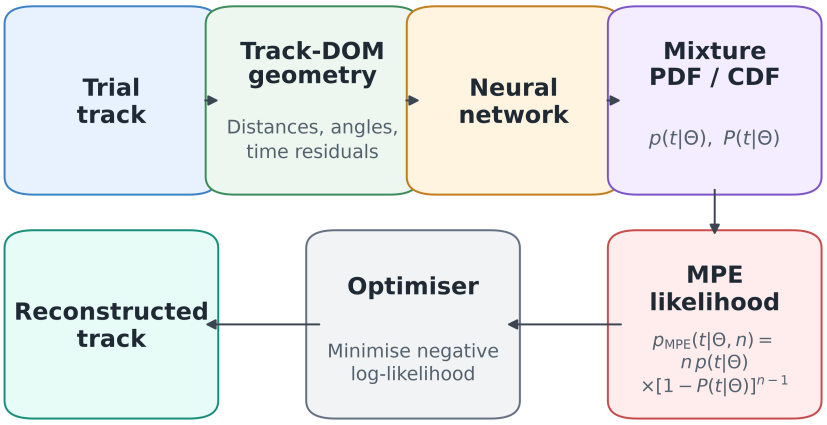


Figure 4.2: TriplePandel reconstruction pipeline. Each trial track defines the track–DOM geometry, the network predicts the analytical timing-distribution parameters, and the MPE likelihood is evaluated before the next minimiser update.

illustrated in Figure 4.2:

1. compute the geometric variables relating the track to every selected DOM;
2. evaluate TriplePandelNET and obtain the mixture parameters;
3. construct the DOM-level PDF, CDF and MPE likelihood terms;
4. sum the negative log-likelihood over the event;
5. update the track parameters through numerical minimisation.

The procedure is repeated until convergence, and the final minimiser output is Θ_{reco}.

The reconstruction uses JAX for array operations, automatic differentiation, just-in-time compilation and vectorisation [38]. These tools support the repeated evaluation of the network and likelihood during optimisation; computational and reproducibility details are collected in Appendix C.

4.6 Infinite-muon and uniform-light hypothesis

The reconstruction considered in this work relies on a simplified description of the muon source. Keeping in mind the physical processes introduced in Chapter 3, the first approximation is to represent the track as an infinite straight line.

TriplePandelMPE models the source as an infinitely long muon with uniform light production along its trajectory, as illustrated in Figure 4.3 [6]. The model has no fitted production vertex, stopping point, or event-specific sequence of localised

energy losses. It only depends on the geometrical relation between this infinite line and each DOM.

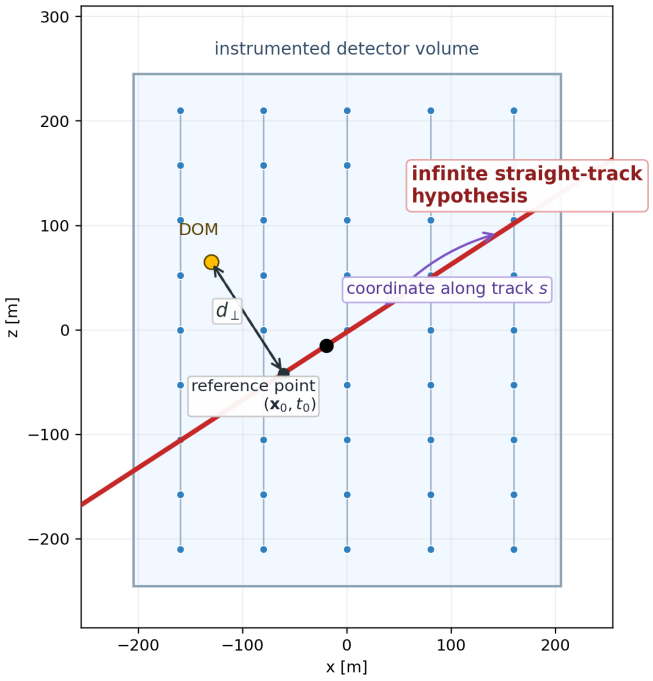


Figure 4.3: Illustration of the infinite muon track hypothesis. The muon is modelled as an infinite straight line, parametrised by a direction and a reference point, while each DOM contributes through its geometry relative to this line.

As discussed in Chapter 3, a physical muon continuously loses energy while propagating through the ice. Its mean stopping power is commonly written as

$$-\left\langle \frac{\mathrm{d}E}{\mathrm{d}x} \right\rangle = a(E) + b(E)E, \tag{4.17}$$

where $a(E)$ mainly describes ionisation losses, while $b(E)E$ contains the contributions from bremsstrahlung, pair production and photonuclear interactions. The relative importance of these radiative processes increases with energy and can lead to large localised energy losses along the muon trajectory [19, 39].

These losses are also important for the detected light. The secondary charged particles produced by ionisation and radiative interactions contribute additional Cherenkov emission, and this contribution becomes increasingly important as the muon energy rises. In the model studied here, however, these individual losses are not represented separately. Their contribution is instead absorbed into a uniform light production along the infinite track.

This distinction is important for the interpretation of the study. The production scans do not attempt to reproduce an event-specific stochastic energy-loss pattern and then reconstruct it with a uniform model. The same uniform infinite-muon description is used consistently in the synthetic generation and reconstruction, so

the neglected localised losses are part of the common model assumption rather than an additional source of model mismatch.

Within this approximation, the overall amount of simulated light is controlled by a dimensionless light-yield scale s_E. It multiplies the expected DOM charge before the Poisson sampling and changes the amount of light observed while leaving the underlying uniform-track description unchanged. The parameter should not be interpreted as an energy reconstructed by the fitter; it is an input to the synthetic event generation. Its reference value and its use in the production scans are defined in Chapter 5.

4.7 Evaluation principle and scope

With these assumptions established, the reconstruction studied in the following chapters can be interpreted within a well-defined domain: an infinite, uniformly emitting muon is generated and reconstructed using TriplePandelMPE under the same detector and optical description. The result asks how well this approximation could perform if its continuous light emission were modelled perfectly within the common forward model. It defines the limit of reconstructions based only on this continuous-emission hypothesis. It is not a limit on a complete event model with localised stochastic losses. Chapter 5 defines this matched-model closure test more precisely, together with the scan strategy, its limitations and the quantities used to compare the generated and reconstructed tracks.

Chapter 5

Methodology

In this chapter, I describe how the simulation strategy used to test the reconstruction model evolved throughout the work. The starting idea was simple: isolate one parameter that defines a track, such as its direction, depth or distance from the detector centre, vary it, and observe how the reconstruction behaves. In practice, the first scans quickly showed that this is only meaningful if the other geometrical effects are controlled as well. A track that is slightly closer to a string, a DOM level or the detector boundary can receive a different amount of useful light and can reconstruct differently even when the parameter of interest has not changed.

The methodology was built progressively. I first used simple and deliberately fixed configurations to understand the response of the model. The structures observed in these early scans were then used to identify possible geometrical biases and to design more general scans in which the nuisance variables are sampled rather than fixed. Some of these exploratory results are kept in this chapter because they explain why the final protocols were chosen.

From this point onward, the study remains within the same controlled framework introduced in Chapter 4. Generation and reconstruction use the same model assumptions, detector geometry and optical description. The aim is not to reproduce every complication of a real IceCube event, but to test how well the reconstruction closes under known and reproducible conditions.

5.1 Principle and scope of the study

The main purpose of this work is to determine the limits of the infinite-muon approximation. More specifically, I ask how well an infinite-muon reconstruction could perform if the continuous light emission and the common detector and optical response were modelled perfectly within the test. This is not a search for one favourable track, nor a test of whether every physical assumption encoded in TriplePandelNET is exact in the real detector. The known generating track provides the controlled reference needed to measure how precisely the model closes, how that precision changes when one condition is varied, and where limitations remain.

For this reason, no recorded IceCube event is used in the study. There is no

atmospheric-neutrino background, detector noise, event selection from real data or real-time processing. The simulated object is not a complete neutrino interaction either. It is the infinite muon track introduced in Chapter 4, described by a reference point, a reference time, a direction and a light-yield scale. The neutrino vertex and hadronic cascade are not simulated. In the same spirit, the photomultiplier-tube amplification, single-photoelectron charge response, complete waveform digitisation and pulse-extraction chain are not reproduced. The generator directly produces one sampled first-pulse time and one integer charge proxy for each hit DOM. The only additional DOM-electronics response applied explicitly at this stage is the timing uncertainty described below.

This distinction matters for the interpretation of everything that follows. The neutrino direction, the direction of the muon produced in the weak interaction, and the reconstructed muon direction are not exactly the same object. The neutrino–muon production angle, missing energy and the position of the neutrino interaction are outside the present study. The quantities reported below compare the generated and reconstructed *muon* tracks.

5.1.1 From a known track to a synthetic event

The first step is to generate an event for which the answer is already known. The geometry file supplies the position and identity of every DOM in the detector, while the parameters of the true muon track are chosen explicitly. From this known track, the forward model predicts the response expected at each DOM. The complete sequence can be summarised as

$$\begin{aligned} \Theta_{\text{true}} &\longrightarrow \{\mu_i,\ p_i(t)\}_{\text{DOMs}} \longrightarrow \{Q_i,\ t_i\}_{\text{hit DOMs}} \\ &\longrightarrow \widehat{\Theta}_{\text{reco}} \longrightarrow \text{reconstruction metrics.} \end{aligned} \tag{5.1}$$

Here, Θ_{true} contains the reference position, reference time, zenith and azimuth selected for the true track. The index i labels a DOM, μ_i is its corresponding unscaled expected detected-charge response, and $p_i(t)$ is the predicted single-photon arrival-time density. The trained response and its calibration already encode the effective probability that light incident on the PMT produces a registered signal above the discriminator; the simulation does not reproduce the PMT charge-amplification process itself. The fixed light-yield scale s_E, introduced in Chapter 4, sets the Poisson expectation $\lambda_i = s_E \mu_i$. The integer sampled charge proxy is then

$$Q_i \sim \text{Poisson}(\lambda_i) = \text{Poisson}(s_E \mu_i)\,. \tag{5.2}$$

In the implemented charge-generation convention, the expected response μ_i is retained only when the perpendicular distance between the true track and DOM i satisfies $d_{\perp,i} < 400$ m. It is set to zero at larger distances before the Poisson draw. This is a fixed convention of the synthetic response, not an event-selection cut and not a geometrical boundary of IceCube. It is used in every production. The corresponding computational step is described in Appendix C.

A positive expected charge does not guarantee that a DOM is present in one particular event: the Poisson draw can still give $Q_i = 0$. In the simulations, a *hit*

DOM is defined as a DOM for which $Q_i > 0$. For every hit DOM, a first-pulse time t_i is drawn from the multiple-photoelectron timing distribution conditional on Q_i, and a 3 ns Gaussian fluctuation is added to represent the timing uncertainty of the DOM electronics. As a consequence, repeating exactly the same true track does not reproduce exactly the same event; the charges, hit pattern and pulse times fluctuate from one realisation to another.

Of course, this synthetic hit definition is simpler than the complete readout used in the real detector. The discriminator efficiency enters only through the calibration of the expected response; no explicit PMT pulse-height distribution, threshold-crossing draw, local-coincidence logic or waveform trigger is simulated. No independent dark, radioactive or detector noise hits are added either. In this study, event-to-event fluctuations come only from the Poisson charge draw, the first-photon sampling and the stated electronics timing uncertainty.

This distinction between an expected response and an actual sampled hit became important very early in the work, especially in the first boundary studies. If a line passes close to one row of strings, the nearby DOMs receive a larger expected charge. The Poisson sampling then changes not only the total observed charge but also the probability that these DOMs are present in the timing fit. This helped explain why moving a track by only a few tens of metres could change the reconstruction significantly. Proximity to the DOMs is not an independent correction applied afterwards; it directly changes the amount and the spatial distribution of information available to the likelihood. The exploratory evidence for this statement is shown in Sections 5.3.1 and 5.3.3; the resulting detector-centred production is reported in Section 6.5 and interpreted in Section 7.6.

Once the synthetic event has been generated, only the detector information is given to the fitter. Generation and reconstruction use the same detector geometry, optical-medium description, expected-charge and timing responses, photon-detection calibration and light-yield convention, but the true track itself is retained only for the final comparison. The fitter receives the DOM positions, sampled first-pulse times and charges. It does not receive Θ_{true}, nor does it simply reuse the PDFs evaluated for the true track.

Instead, for every trial hypothesis, TriplePandelNET evaluates a new set of timing distributions corresponding to that candidate track. The negative log-likelihood then measures how well those distributions describe the sampled pulses, and the reconstructed track is the hypothesis that minimises this objective. This is an important point for the rest of the chapter: the PDFs are used to generate the synthetic event and are recalculated during the fit, but the PDFs themselves are not what is being reconstructed. The fitted object is still the muon track.

The event-generation and likelihood sequence is described in Appendix C. The likelihood and track parametrisation were introduced in Chapter 4 and are not rederived here.

5.1.2 Reference light yield and statistical realisations

For all simulations in this study, I use the common light-yield scale

$$s_E = 1.425. \tag{5.3}$$

Chapter 4 describes this scale in more detail. It accounts, on average, for the additional Cherenkov light produced by secondary particles. The infinite-muon light production nevertheless remains uniform [40]. The value used here is motivated by the parametrisation of Rädel and Wiebusch [40], who showed that low-energy losses below approximately 0.5 GeV can be treated as a quasi-continuous contribution along the muon track. For a nominal muon energy of 100 TeV, this contribution corresponds to an additional Cherenkov light yield of about 42.5% relative to the reference muon, which gives the factor $s_E = 1.425$ used throughout this work.

This does not mean that a complete 100 TeV muon is explicitly propagated through the detector. Instead, the expected light yield of the uniform infinite track is scaled to reproduce the average continuous contribution associated with such a high-energy muon. Event-specific localised stochastic losses are not generated individually, and s_E is not an energy parameter reconstructed by the fit. It is a light-yield convention used to define the synthetic detector response within the approximation considered in this study.

A single realisation is of course not enough to characterise one configuration. Throughout this thesis, a *light realisation* means one independent draw of the Poisson charges and first-pulse times for the same expected response and track geometry. It changes neither the optical model nor the true line. Each bin contains many stochastic events, with sampled positions and light realisations. Their random-number seeds are recorded so that the samples remain reproducible. Averaging over these realisations reduces the influence of one unusually favourable or unfavourable light pattern. What it does not remove, however, is the geometrical specificity of a line that is kept fixed relative to the detector. This difference became one of the main reasons for changing the scan strategy later in the work.

For every event, the reconstruction is started from several nearby support-point hypotheses together with a reference-time prescan. These are different initialisations of the same fit, not independent simulated events. The initial direction and position are deliberately kept close to truth because the question studied here is the local performance of the reconstruction once the fit is already in the neighbourhood of the correct solution.

The setup is a *local matched-model closure test*, rather than a complete blind reconstruction from an arbitrary first estimate. "Matched model" means that generation and reconstruction share the same forward model and conventions. "Closure" means that this internally consistent chain can be compared with the truth that generated it, and "local" specifies that the optimiser is tested near the true solution rather than at the global seed-finding stage. The computational workflow is documented in Appendix C.

5.1.3 Meaning of the theoretical limit

With this setup, the expression *theoretical limit* has a specific meaning in this thesis. Here, it refers to the reconstruction performance that remains when generation and fitting are kept inside the same controlled and matched framework. Even in this favourable case, the result is not expected to be perfect: finite and Poisson-sampled light, the sparse and discrete DOM geometry, weakly constrained degrees of freedom of an infinite track, and the numerical optimisation can still produce a residual error.

This definition is intentionally narrower than the experimental angular resolution of IceCube. It is an empirical matched-model benchmark defined by the stated simulation and fitting protocol, not a general statistical bound derived analytically from the likelihood. Real detector noise, calibration uncertainty, generator–reconstruction ice mismatch, realistic stochastic losses, the finite start and end of a muon, the neutrino–muon production angle and real-event selection are all excluded. The question is: if generation and reconstruction share the same model, how well can the reconstruction close on its own truth? For the same reason, a good closure result cannot demonstrate that the encoded ice description is correct in nature. If a feature is shared by both generation and reconstruction, this test cannot independently validate it.

The scans can nevertheless recover structures associated with the detector geometry and with the depth-dependent optical profile. This is used only as a qualitative consistency check that the trained network is being evaluated in the domain for which it was developed, not as a validation against real IceCube data. This distinction also makes the benchmark useful for deciding where further modelling effort should go. If substantial limitations remain when the continuous emission is modelled consistently, improving that component or its geometrical treatment may still be valuable. If closure is already precise and the dominant degradation appears only after more realistic effects are introduced, effort is better directed towards stochastic energy losses, calibration mismatches, detector noise or global seeding. The result is thus a limit on what can be achieved using only the continuous light emission of the infinite-muon approximation.

The full geometry contains all 5160 DOMs on 86 strings, including the additional DeepCore strings. These positions were also present during the model training and contribute inseparably to both the generated hit patterns and the reconstruction. The simulated modules do not carry a separate HQE response. Consequently, when a track crosses the DeepCore region, the effect of the denser geometry present in the model can be studied, but this setup cannot isolate an HQE response or switch DeepCore on and off as an independent reconstruction variable.

5.2 Observables used to evaluate the reconstruction

Once an event has been reconstructed, I still need a good way to decide whether the result is actually satisfactory. Angular error is the most important observable for this work, but using it alone can be misleading. Two infinite lines can be perfectly parallel and have zero angular separation while remaining several metres apart, as

illustrated in Figure 5.1. The opposite situation can also occur: two non-parallel infinite lines may cross outside the instrumented volume and have zero closest-line distance even though their directions are different. In addition, an excellent error among the successfully reconstructed events would not be very informative if a large fraction of the requested events had failed. For this reason, I use several complementary observables to describe direction, spatial agreement, available light, distribution tails and event survival.

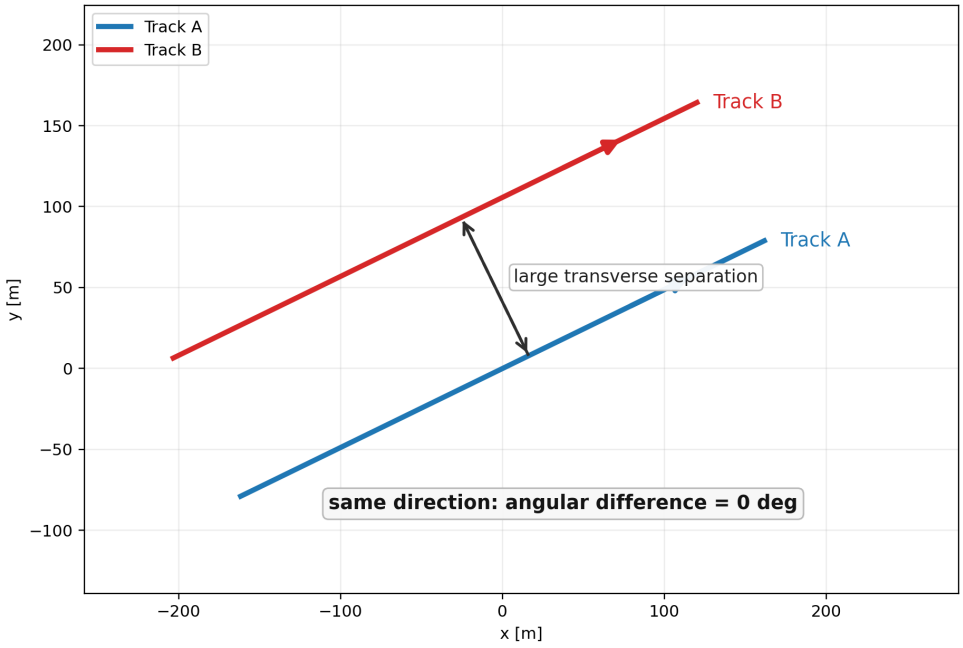


Figure 5.1: Two parallel tracks have no angular separation but retain a finite spatial distance. Directional and spatial metrics are complementary.

The three-dimensional angular error is

$$\Delta\psi = \arccos[\mathrm{clip}(\hat{\mathbf{u}}_{\mathrm{true}} \cdot \hat{\mathbf{u}}_{\mathrm{reco}}, -1, 1)]\,. \tag{5.4}$$

Here, $\hat{\mathbf{u}}_{\mathrm{true}}$ and $\hat{\mathbf{u}}_{\mathrm{reco}}$ are the generated and reconstructed unit direction vectors, and $\mathrm{clip}(x, -1, 1)$ restricts numerical rounding excursions of their dot product to the domain of the arccosine. The angular error $\Delta\psi$ is non-negative and describes the magnitude of the directional discrepancy. It is evaluated event by event, so its distribution describes resolution but does not show whether the reconstructed directions share a common displacement.

Signed coordinate residuals are retained as simple diagnostics. For event i, I denote them by

$$\begin{aligned} \Delta\theta_i &= \theta_{\mathrm{reco},i} - \theta_{\mathrm{true},i}, \\ \Delta\phi_i &= \mathrm{wrap}_{[-\pi,\pi)}(\phi_{\mathrm{reco},i} - \phi_{\mathrm{true},i})\,. \end{aligned} \tag{5.5}$$

The azimuth difference is wrapped across the $0°/360°$ boundary. Bin-level means are written as $\langle\Delta\theta\rangle$ and $\langle\Delta\phi\rangle$. These separate residuals are useful when the coordinates are well conditioned, but they do not form a coordinate-independent angular bias on the sphere. The symbols $b_{3\mathrm{D}}$, b_θ and b_ϕ are reserved below for the aligned directional bias and its local components.

For the final azimuth, zenith and detector-centred radial scans, the coherent directional bias is evaluated in three dimensions. For each event i, the rotation

$$R_i = R_y(-\theta_{\mathrm{true},i})R_z(-\phi_{\mathrm{true},i}) \tag{5.6}$$

maps the true direction onto the positive local z axis. The same rotation is applied to the reconstructed direction. The rotated reconstructed vectors can then be averaged inside a bin because their true directions are aligned. The opening angle between the normalised mean vector and the positive local z axis defines the aligned three-dimensional bias magnitude b_{3D}. Its tangent-plane components are denoted by b_θ and b_ϕ, where b_ϕ is a physical local azimuthal-arc component rather than a raw difference in ϕ. By construction, $b_{\mathrm{3D}} = \sqrt{b_\theta^2 + b_\phi^2}$. The complete definition is given in Section A.1.

Since only a rotation is applied, the opening angle of each event is preserved and the calculation remains valid close to the zenith poles. At the exact poles, however, the signed orientation of the two local components depends on the chosen roll convention. The magnitude b_{3D} is the primary bias quantity there, while the components are shown only as diagnostic information.

Spatial quantities are useful as supporting diagnostics, although they are not kept as a fixed panel in every final result. The *line distance* is the minimum separation between the two complete infinite lines. Because it is unchanged when a support point is shifted along its own line, it is the more relevant quantity when the question is whether the reconstructed trajectory crosses the same detector region as the true one. The *support-point distance*, on the other hand, is simply the Euclidean distance between the true and reconstructed reference points. These points only parametrise the two lines; they are not physical neutrino-interaction vertices. This quantity is removed from the standard final plots, but retained when it helps to understand an exploratory diagnostic. The formulae, parallel limit and auxiliary transverse decomposition are given in Section A.1.

The reconstruction errors alone also do not tell us how much information was available to the fit. For that reason, I keep track of the total sampled charge and of the number of hit DOMs,

$$Q_{\mathrm{tot}} = \sum_i Q_i, \qquad N_{\mathrm{hit}} = \sum_i \mathbf{1}(Q_i > 0). \tag{5.7}$$

These are not reconstruction errors themselves. Their role is to show whether a change in resolution occurs at the same time as a change in the amount or spatial multiplicity of the observed light. This became particularly useful in the radial and local-geometry studies, where moving the track can directly change which DOMs contribute to the likelihood.

Here, Q_{tot} is the sum of the sampled charge proxies over all DOMs, N_{hit} is the number of DOMs with a positive sampled charge, and $\mathbf{1}(\cdot)$ is the indicator function, equal to one when its condition is true and zero otherwise.

Let N_{gen} denote the number of requested events, N_{det} the number for which at least one DOM satisfies $N_{\mathrm{hit}} > 0$, and N_{reco} the number of detected events for which

the fit returns a storable reconstructed track. The event-survival quantities are

$$f_{\mathrm{det}} = \frac{N_{\mathrm{det}}}{N_{\mathrm{gen}}}, \qquad f_{\mathrm{reco}} = \frac{N_{\mathrm{reco}}}{N_{\mathrm{gen}}}, \qquad f_{\mathrm{reco|det}} = \frac{N_{\mathrm{reco}}}{N_{\mathrm{det}}}. \tag{5.8}$$

The detection fraction f_{det} is the fraction of requested events that contain at least one hit, f_{reco} is the fraction of all requested events that are reconstructed, and $f_{\mathrm{reco|det}}$ is the conditional reconstruction fraction among detected events. If an event is not reconstructed, its reconstruction metrics remain undefined, but the event still remains in the denominator of the relevant success fractions. This convention is kept deliberately: otherwise a scan could appear to have excellent resolution simply because only a small favourable subset of events survived. The hit-DOM condition and the exact detected and reconstructed populations entering these fractions are specified in Section A.3.

Another point that became important when looking at the first distributions is that one number cannot describe the full event population. The error distributions are generally non-Gaussian and can contain long tails. The median p_{50} is used for the typical successful event, together with p_{68} and p_{95} to show how the poorer reconstructions progressively enter the distribution. A directional-bias estimate must also be interpreted differently from a positive event-wise error: the bias describes the displacement of the aggregate direction, whereas the positive error describes the magnitude of each event's discrepancy.

For the aligned directional bias and the signed coordinate diagnostics, we also need to check whether an apparent shift remains stable despite the finite number of events. The final results use bootstrap resampling for this purpose. Within each bin, complete events are drawn with replacement and the bin-level statistic is recomputed. For the aligned bias, the rotated vectors are resampled and their mean direction is recalculated in every replicate. The resulting interval estimates the sampling uncertainty of the bin-level bias; it is not the same quantity as the p_{68} or p_{95} spread of individual events and it does not remove or correct the bias. The bootstrap formula, output columns, units and evaluation populations are collected in Appendix A.

The standard detector-wide production estimate uses 1,000 event-level resamples in every bin. The local DOM coordinate-residual summaries use a different unit because their target is an average over detector locations. The signed mean is first evaluated for each DOM, and the mean over the selected DOM population is assigned an uncertainty from 2,000 DOM-level resamples. This construction is used for both the primary 24-DOM curves and the later comparison between the full 42-DOM sample and its centre + intermediate subset. The two interval types answer different questions and are not interchangeable.

This is also why the exploratory figures below do not all contain exactly the same panels or use the same normalisation. At that stage of the work, each scan was designed to answer a specific question. If moving a track changes its illumination, charge and hit multiplicity are essential; when a directional shift is suspected, a signed angular residual is more informative. I keep this case-by-case presentation because it reflects how the methodology was developed. For the final results, however, a common core is used, consisting of angular-error percentiles, the relevant

directional-bias estimate, total charge, hit-DOM multiplicity and event success fractions. Bootstrap confidence intervals are added to the bias estimates, while spatial quantities are kept only when they help answer a particular geometrical question.

5.3 From fixed tracks to controlled ensembles

The next question is how to vary the track parameters without mixing several effects at the same time. An event depends on azimuth, zenith, depth, position, the fixed light-yield convention, distance to the instrumented region and the relation of the line to the nearest DOMs. My first approach was the most direct one: keep everything fixed and vary only one parameter. The idea itself remains the basis of the final scans, but the first results showed that "keeping the other parameters fixed" was not always the best way to isolate the quantity I wanted to study.

The reason is the discrete geometry of IceCube. The detector is not a continuous cylinder: its strings form an approximately hexagonal but irregular footprint, the standard DOMs are separated vertically by about 17 m, and DeepCore introduces a denser region (Figure 3.1b). Repeating one fixed line also repeats its particular distance to the closest strings and DOM levels. In several early tests, this local placement affected the reconstruction more strongly than the parameter that was supposed to be isolated.

From these observations, I changed the strategy. Instead of representing one bin by one fixed line, the final scans keep one main variable binned and sample the nuisance variables from the same declared distribution in every bin. This does not remove the discrete geometry of the detector—nor should it—but it prevents one accidentally favourable or unfavourable placement from defining an entire curve.

The exploratory scans below are kept when they explain how this conclusion was reached. They are not all final physics results. For me, a diagnostic scan asks why one particular configuration behaves as it does; a production scan asks whether the same dependence remains once the nuisance geometry has been sampled according to a reproducible protocol.

5.3.1 Azimuth, corridor geometry and transverse position

I started the azimuth study with horizontal tracks passing through the detector centre at regularly spaced directions. Horizontal tracks were useful because they remain at one nominal depth, while the central position allowed a large part of the detector to contribute. At this stage, the same geometrical line was simply repeated many times at each angle in order to average the stochastic charge and timing fluctuations. One of the largest angular errors appeared near 70°, as shown in Figure 5.2. At first, this looked like a possible azimuth dependence of the reconstruction.

Looking at the detector from above changed that interpretation. The 70° line passed through a corridor between two rows of strings rather than close to one of them (Figure 5.3). I kept the direction and depth fixed and translated the whole line across this corridor, as illustrated in Figure 5.4. If the original peak was really

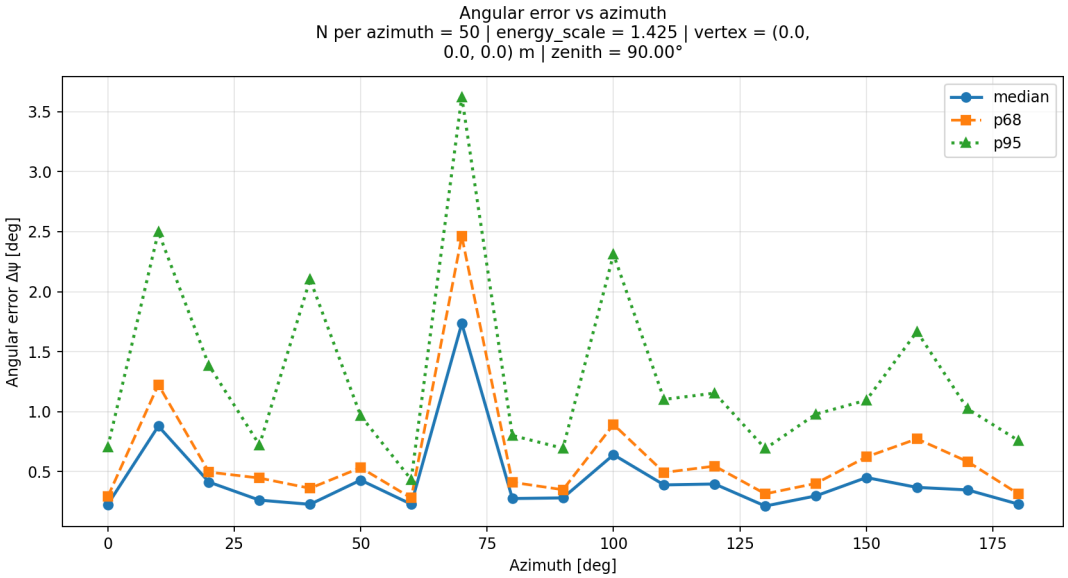


Figure 5.2: Angular-error summary from the first azimuth scan. Repeating stochastic pulses reduced the fluctuations of one line but did not change its position relative to the strings. The feature near 70° led to the corridor diagnostic.

an azimuth effect, moving the same direction sideways should not remove it. If it was mainly a local-placement effect, the reconstruction should change as the line approached the neighbouring string rows. The final diagnostic used nine offsets between approximately −50 and +50 m, with 100 stochastic realisations at each offset. The corresponding angular, spatial and charge variations are compared in Figure 5.5.

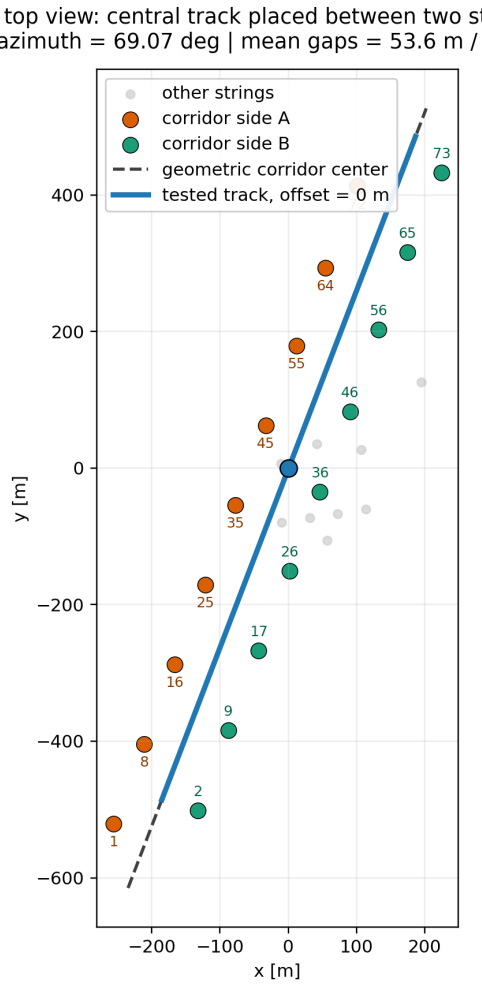
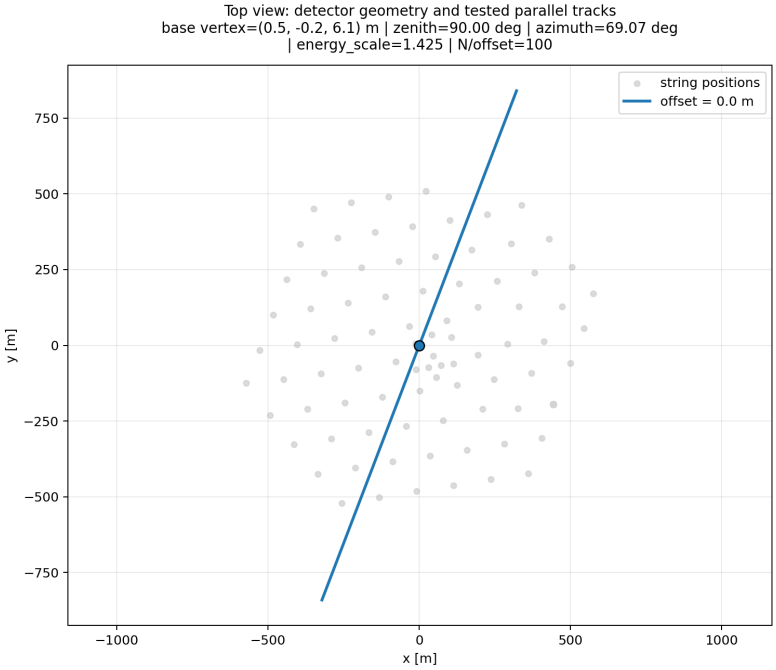
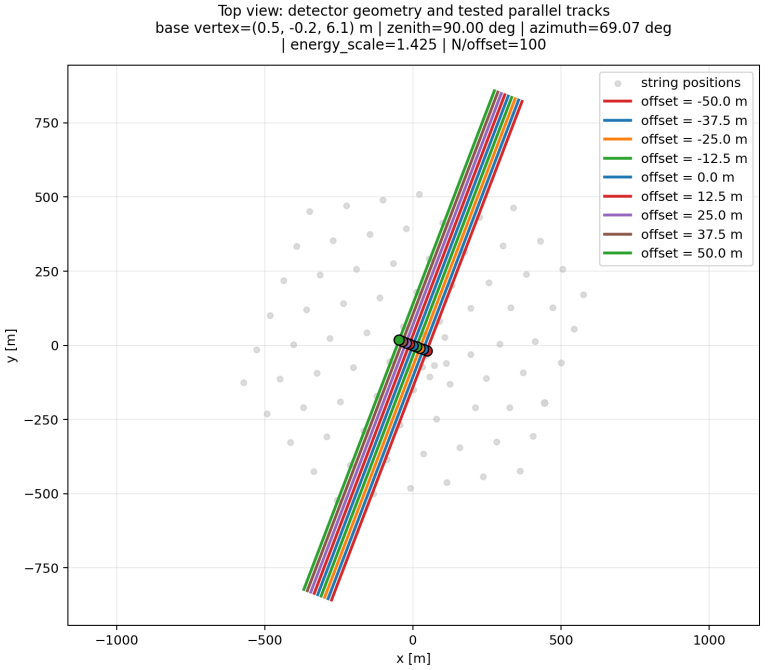


Figure 5.3: Top view of the fixed line at the azimuth where the first scan showed its largest feature. The line passes between two neighbouring rows of strings rather than close to one row, which motivated the transverse translation test.

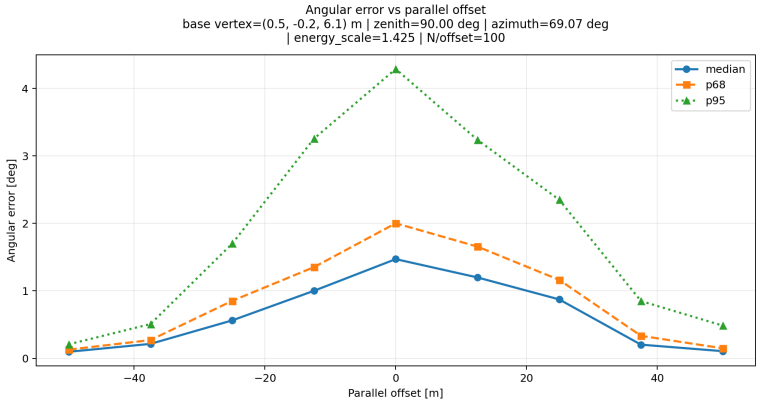


(a) Original fixed line at zero offset.

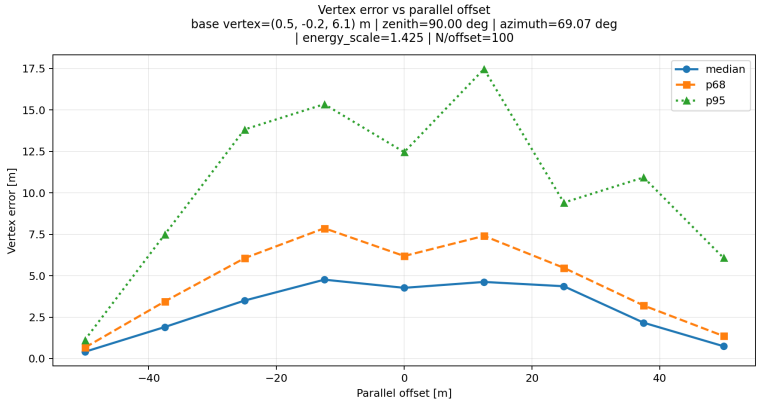


(b) Family of translated parallel lines.

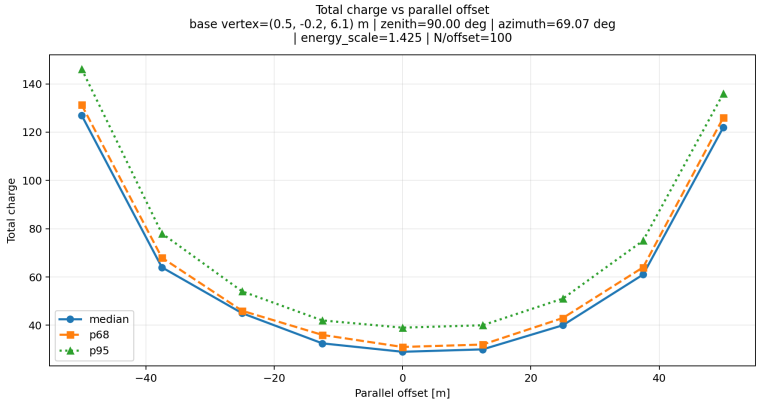
Figure 5.4: Geometry of the transverse-offset diagnostic. The direction and depth remain fixed while the entire line is translated across the corridor. The two panels separate the original placement from the full family of offsets used in the comparison.



(a) Angular reconstruction error.



(b) Support-point distance.



(c) Total sampled charge.

Figure 5.5: Effect of translating the same true direction across the corridor. Directional and spatial errors are largest around the central placement, where the sampled charge is lowest, and decrease when the line approaches a string row. The comparison demonstrates a local placement dependence; it does not establish that this azimuth is globally unfavourable.

This test was important because it changed the meaning of the first azimuth curve. One hundred Poisson and timing realisations are a good statistical ensemble for one geometrical line, but they are still one line relative to the detector. Repeating it does not average the corridor away. From what I observed here, the same idea could be applied directly to the azimuth scan: instead of using one central line per angle, I sample the true direction within each 5° bin and translate the line over a wide transverse band. The geometry is illustrated in Figure 5.6.

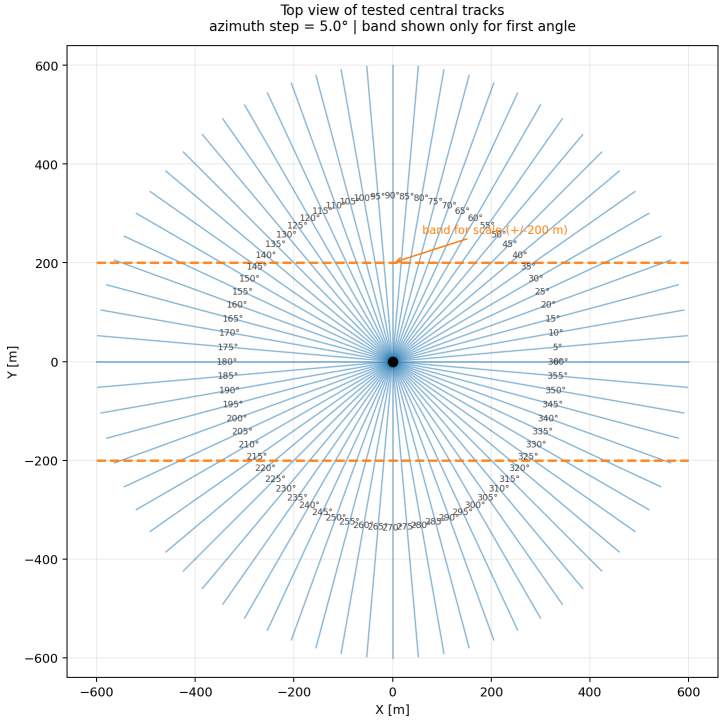


Figure 5.6: Top view of the broadened azimuth geometry. Track directions cover the full angular range while their perpendicular position is sampled inside the same transverse band in every azimuth bin. This replaces the single central line used in the first diagnostic.

The 400 m-wide version then revealed a second effect. The growing separation between the median, 68th and 95th percentiles initially looked like it could come from a common offset or bias (Figure 5.7). However, the event-level histogram did not support that interpretation (Figure 5.8). When I plotted the angular error against the position inside the band, the origin of the tail became clearer: the error increases with the absolute transverse offset (Figure 5.9). Tracks near the outer parts of the band are less contained and are observed by a different subset of strings. Since the depth is fixed, this is mainly a geometrical and containment effect rather than a change of ice layer.

The broad band is nevertheless kept in the production scan. The point is not to make every event equally well contained, but to apply the same spatial distribution in every azimuth bin. The resulting curve represents an azimuth-averaged reconstruction performance over that declared region instead of the performance of one specially chosen central line.

Some exploratory curves also showed approximate structures separated by about

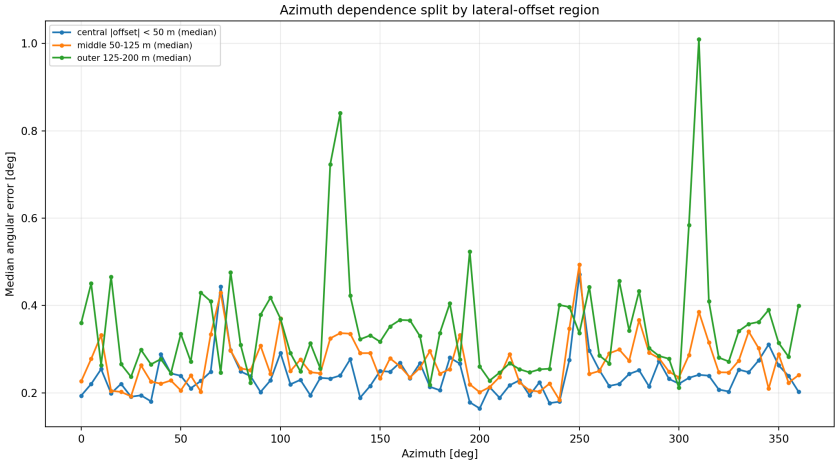


Figure 5.7: Azimuth curves separated by transverse-offset region. Diagnostic used to test whether the separation of the global percentiles represented one common bias.

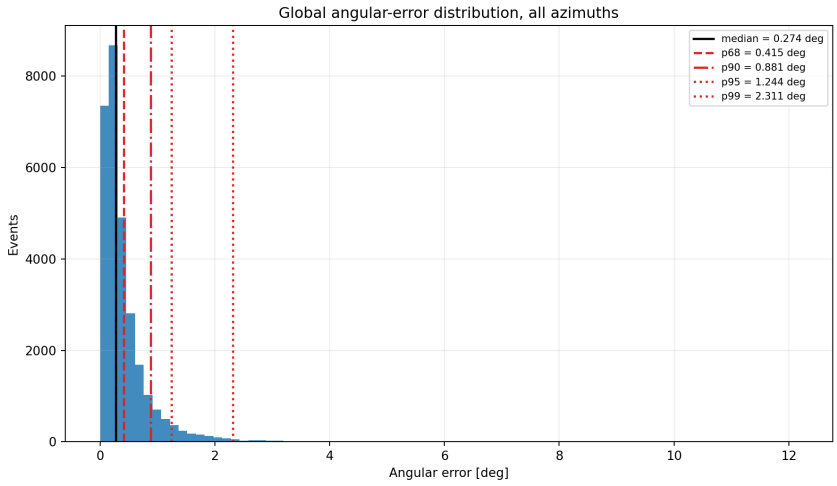


Figure 5.8: Event-level angular-error distribution over all azimuths. Diagnostic used in addition to Figure 5.7. The offset-resolved curves and the strongly asymmetric event-level distribution instead point to a position-dependent tail.

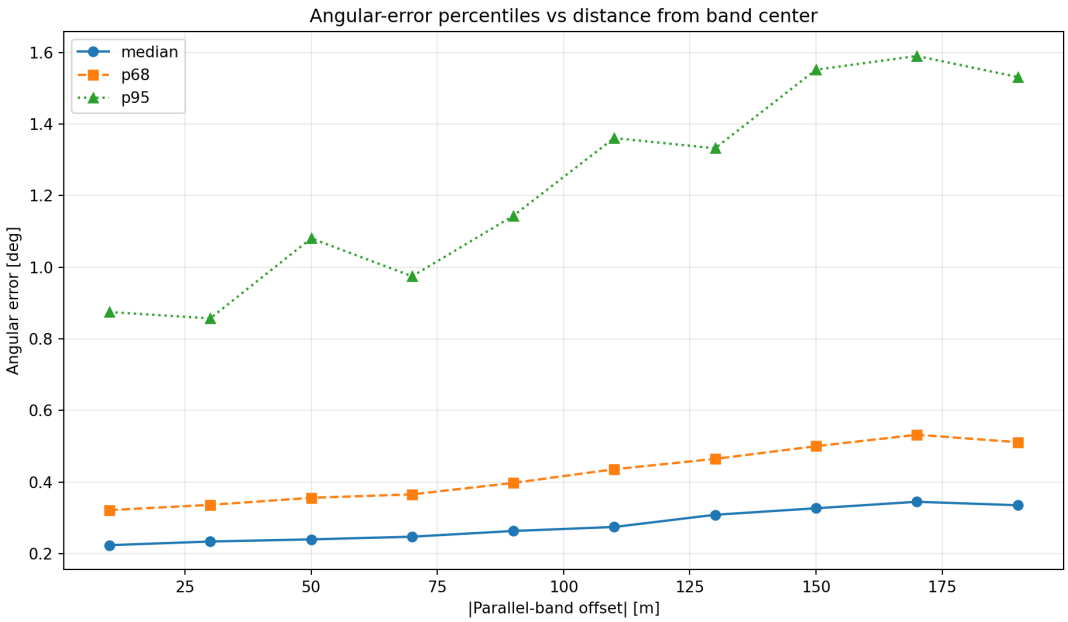


Figure 5.9: Angular-error percentiles as a function of the absolute position inside the transverse band. The deterioration towards the outer part of the sampled region motivated treating containment and local placement as nuisance variables of the azimuth scan.

60°. This is suggestive because the string layout has an approximately six-fold structure, but the exploratory scan alone is not used to assign a cause. The final comparison at three depths, together with the charge and hit-DOM information, is needed before interpreting any residual azimuthal modulation.

5.3.2 Depth and the vertical DOM spacing

Depth needs to be treated separately because the detector and the optical medium are not vertically symmetric. Absorption and scattering change with the impurity content of the ice, including the dust-rich interval and the clearer ice below it [36]. For that reason, a radial or purely geometrical scan cannot replace a dedicated depth study.

My first depth scans were again very direct: I translated the same horizontal track vertically, first in 50 m steps and later with a finer spacing. This was useful to see whether any structure appeared, but it kept the same azimuth and repeatedly selected particular distances to the DOM levels. The resulting curve could mix the true depth dependence with a corridor effect and with a vertical alignment effect. Later versions sampled the azimuth over the full circle so that no single horizontal direction defined a complete depth bin.

There was a second issue with using exact depth values. If I simulate only one value every 10, 25 or 50 m, all events in that bin can sit at a special distance from the nearest DOM layer. Choosing steps of exactly 17 m would not fix this either, because it would simply preserve the same relative offset from one layer to the next. The real DOM coordinates are also not a set of perfectly coplanar rings. The final

solution is to use layers whose width matches the standard 17 m DOM spacing and sample the true z position inside each layer. The bin still identifies the depth region that is being tested, while the within-layer randomisation averages over the accidental proximity to a particular DOM row. Figure 5.10 illustrates this idea.

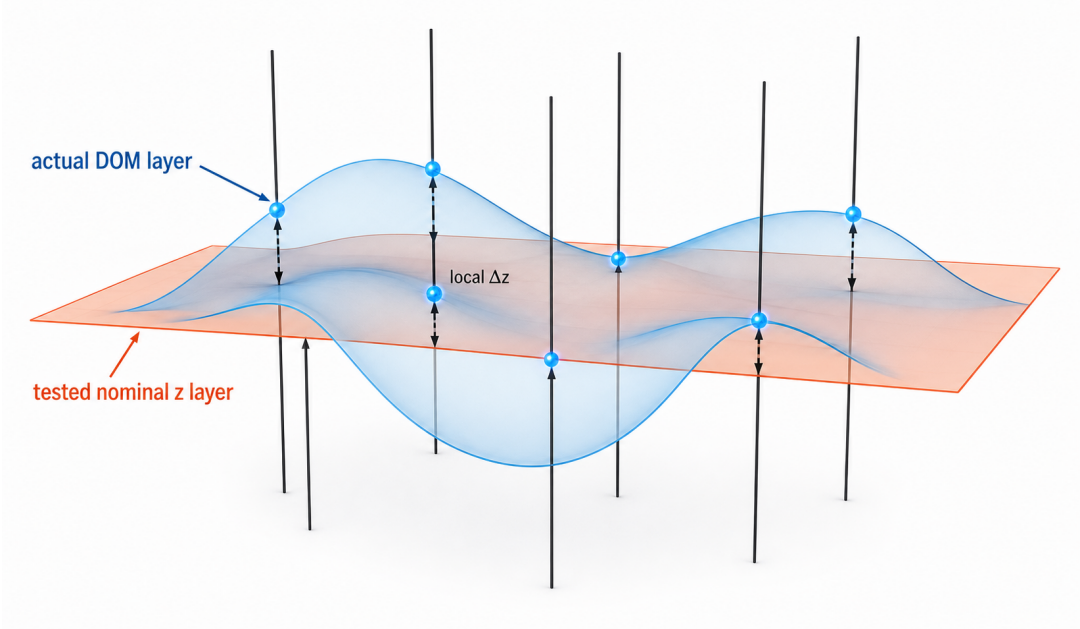


Figure 5.10: Illustration of the vertical sampling used inside one depth layer. The nominal layer identifies the tested depth region, while the event position is displaced relative to the nearby, non-perfectly coplanar DOM levels. Sampling within the layer avoids repeating one special vertical alignment throughout the scan.

5.3.3 From selected detector sides to a radial coordinate

I initially approached the detector boundary in a similarly intuitive way. I selected several visible sides of the detector and placed horizontal tracks parallel to them, as shown in Figure 5.11. The tracks were moved from the centre towards each side using fractions of the available distance, initially in steps of about 15%. This construction was easy to visualise, but the first comparisons showed its limitation: the detector footprint is irregular, the different sides are not equally far from the centre, and the same fractional position can place a track much closer to a string on one side than on another. In other words, the coordinate I was varying was not physically equivalent from one side to the next.

I then tried using successive rows of DOMs approximately parallel to the selected sides. The comparison became cleaner, but the same problem did not disappear completely. The response still differed from one side to another, especially when one of the tested lines happened to pass close to a DOM row. Figures 5.12 and 5.13 show that the side-dependent angular error was accompanied by a side-dependent amount of observed light. This was a useful result in itself: it confirmed that expected charge, Poisson sampling and local hit multiplicity can strongly affect the reconstruction. At the same time, it showed that this six-side construction was not yet a clean detector-wide measure of the boundary.

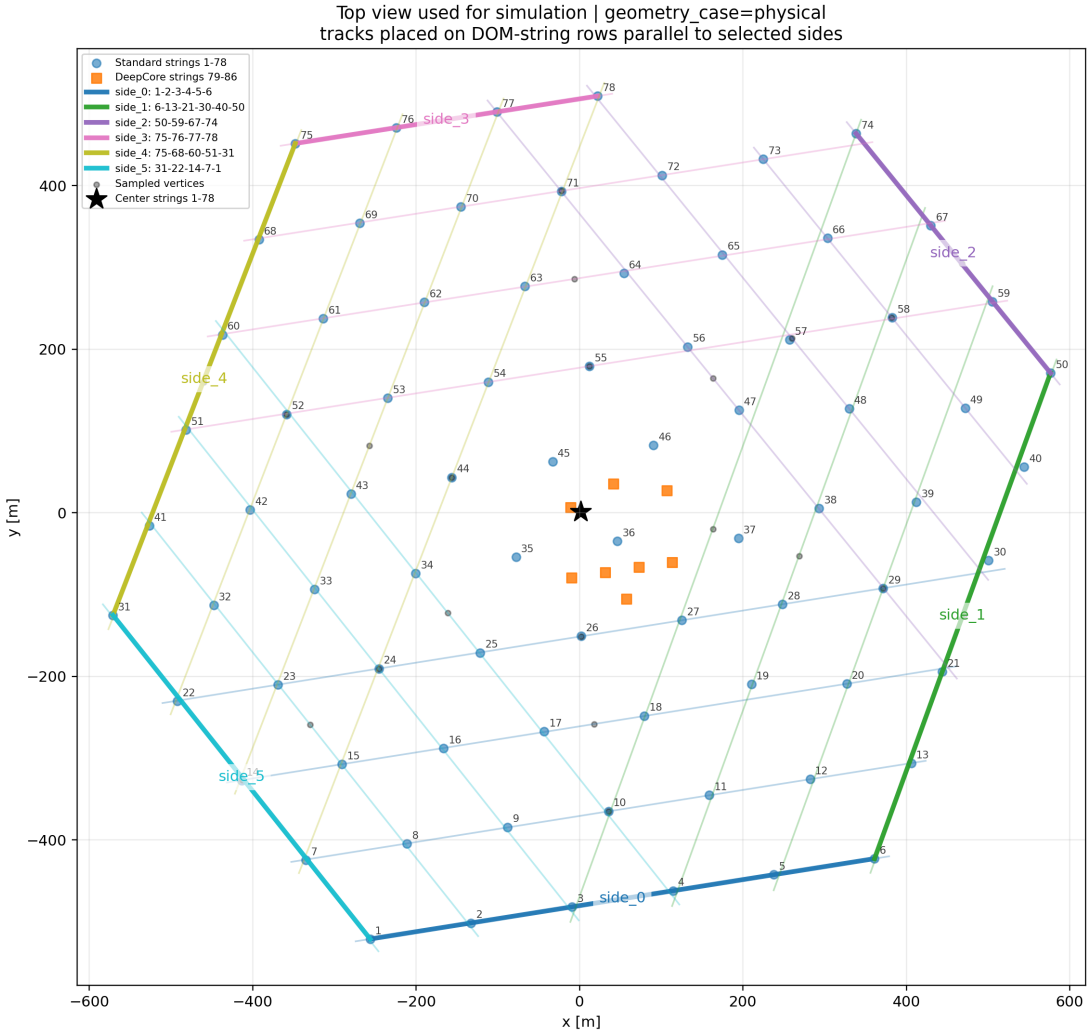


Figure 5.11: Top view of the first boundary construction. For each manually selected side, horizontal tracks were placed along rows approximately parallel to the nearby detector boundary. The six coloured outlines identify the side-specific directions used in this exploratory comparison.

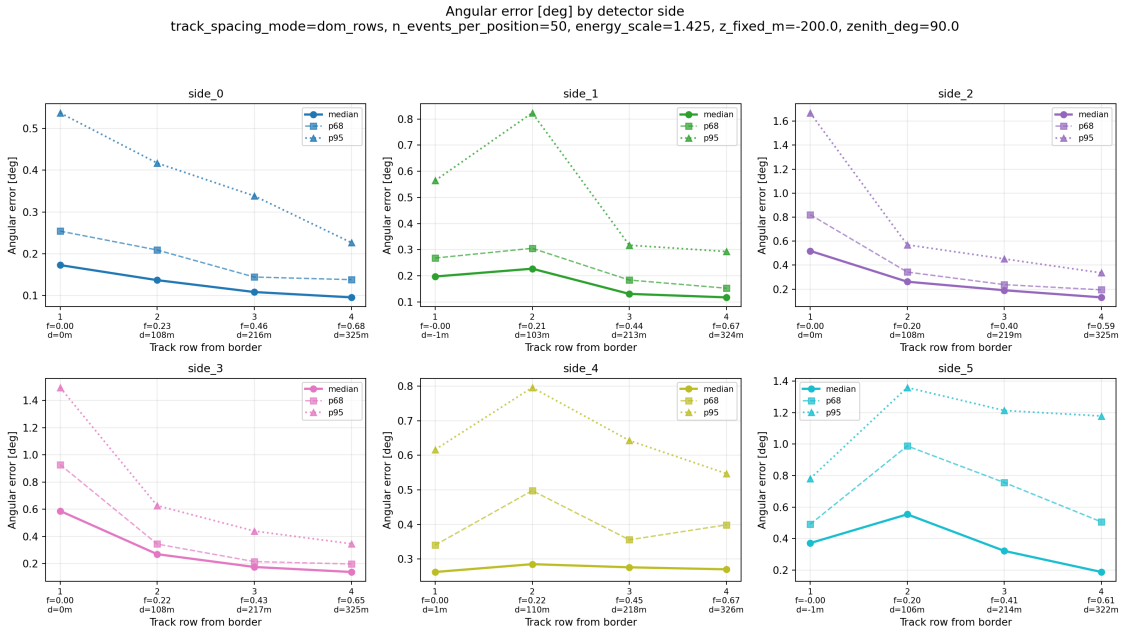


Figure 5.12: Angular-error percentiles for the successive parallel track rows on each of the six selected detector sides. The panels are kept separate because an equal fractional step does not correspond to an equal local distance from the nearest strings on every side.

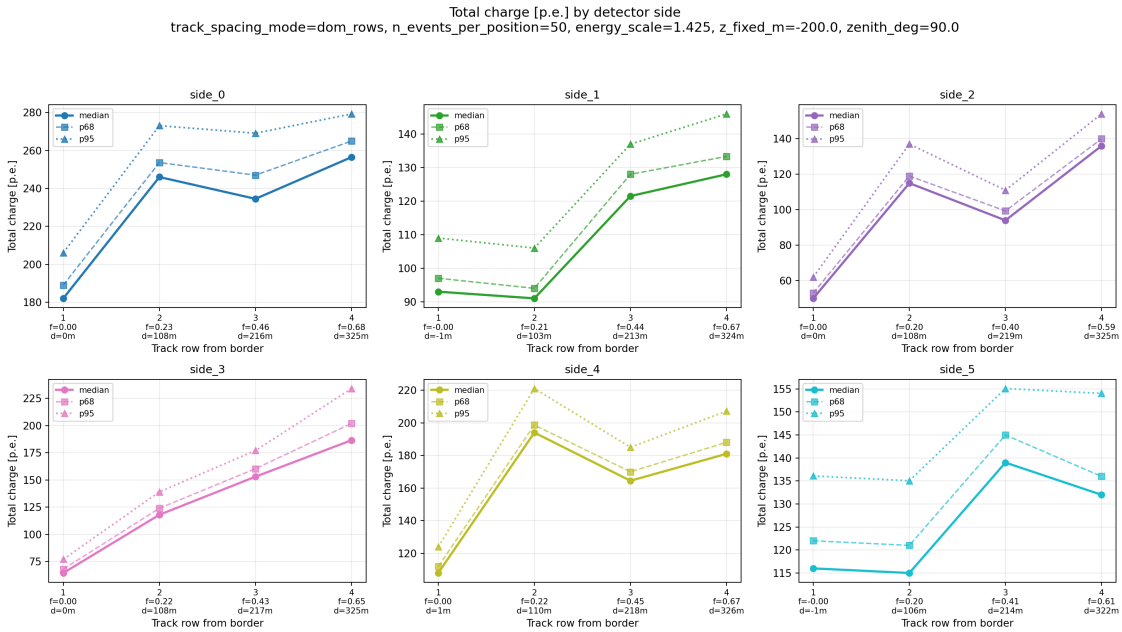


Figure 5.13: Total sampled charge for the same six-side track families as in Figure 5.12. The non-equivalent charge evolution between sides confirms that the construction mixes boundary distance with local proximity to the discrete string rows.

From these boundary tests, I moved to a coordinate that does not depend on choosing one particular detector side. The production scan asks how the reconstruction changes as the closest approach of the infinite line moves progressively away from a chosen detector centre. If $\hat{\mathbf{u}}$ is an isotropic track direction and $\hat{\mathbf{v}}$ is a random unit vector in the plane perpendicular to it, the true line is

$$\mathbf{r}(s) = \mathbf{r}_{\mathrm{c}} + R\hat{\mathbf{v}} + s\hat{\mathbf{u}}, \qquad \hat{\mathbf{u}} \cdot \hat{\mathbf{v}} = 0. \tag{5.9}$$

Here, $\mathbf{r}_{\mathrm{c}}$ is the chosen detector centre, R is the closest-approach radius and s is the signed distance along the line. The shortest distance from the line to $\mathbf{r}_{\mathrm{c}}$ is then R. For each value of R, sampling both the direction and the perpendicular orientation produces an ensemble of tangent lines around the centre in three dimensions. This does not mean that IceCube is assumed to be spherical, and R should not be interpreted as the literal distance to one physical border. Rather, it gives a single reproducible coordinate for following the reconstruction from the central region towards, and eventually beyond, the instrumented detector. This is the reason the radial construction replaced the earlier six-side comparison in the final production.

5.4 Production scans

Once these geometrical effects had been identified, I used the same general rule for the final production campaigns: one variable defines the bins, while the other coordinates are sampled from a distribution that is kept identical from one bin to the next. The aim is not to remove the real detector geometry, but to make sure that a change in the final curve is not caused simply by changing the way the nuisance geometry was sampled.

The detector coordinate system is used throughout. In the supplied geometry, $z = 0$ lies near the vertical centre of the instrumented array, approximately 1950 m below the surface, and positive z points upward. Whenever a scan requires a three-dimensional detector centre, I define the full coordinate explicitly for that protocol rather than inferring it from the surface depth.

The individual production protocols are given below and summarised together in Table 5.1.

Scan	Main coordinate	Events per bin	Total requested	Sampling and fixed conditions
Azimuth	72 bins of 5° at three nominal depths	400 per depth	86,400	Horizontal tracks ($\theta = 90°$); true azimuth sampled inside each bin; transverse offset in a ±200 m band; z sampled in 17 m layers centred at $+50, -100, -250$ m
Depth	59 layers from −500 to +486 m	200	11,800	Horizontal tracks ($\theta = 90°$); 17 m layers; z sampled inside the layer; azimuth sampled in balanced strata over 360°; lateral radius 200 m
Zenith, nominal	37 bins from 0° to 180°	400	14,800	Azimuth uniform; position sampled in a perpendicular disk of radius 225 m around $(0, 0, -200)$ m
Zenith, centred	37 bins from 0° to 180°	400	14,800	Azimuth uniform; 200 m perpendicular disk around $(5.87, -2.52, -23.91)$ m; independent seed
Zenith, upper	37 bins from 0° to 180°	400	14,800	Matched 225 m perpendicular disk around $(0, 0, +200)$ m
Radial	15 closest-approach radii from 0 to 700 m	1,000	15,000	Isotropic directions; tangent position sampled around the selected detector centre; complete detector geometry

Table 5.1: Final detector-wide runs, requested event counts and controlled sampling.

5.4.1 Azimuth

For the final azimuth production, I keep the broad spatial sampling motivated by the corridor study and cover the complete 0°–360° interval. A full turn is required because the track has an orientation: reversing its direction changes the order of the predicted pulse times even if the geometrical line itself is unchanged. The true azimuth is sampled inside each 5° bin rather than fixed at the bin centre, and all tracks have the horizontal zenith $\theta = 90°$. They are translated inside the same ±200 m transverse band for every bin.

I also repeat the production at three nominal depths, above the dust-rich interval, inside it and in deeper clearer ice. Their centres are $z = +50$, -100 and -250 m. As in the dedicated depth study, the true position is not fixed exactly at these values: it is sampled inside the corresponding 17 m layer. This prevents one particular vertical alignment with the DOM levels from defining an entire azimuth curve. Comparing the three productions then shows whether an azimuthal structure persists in regions with different depth-dependent optical properties. The scan still tests reconstruction performance within the prescribed model; it is not an independent measurement of the ice.

5.4.2 Depth

The final depth scan follows directly from the problems identified in the early vertical translations. I keep the tracks horizontal ($\theta = 90°$) so that each event remains within one nominal depth region, but I no longer represent that region by one exact line. The 59 layers cover $z = -500$ to $+486$ m. Within every layer, z is sampled rather than fixed, the azimuth is sampled in balanced strata over the full circle, and the lateral position is randomised within the selected radius. Here, *stratified* means that the full 360° range is divided into angular strata represented with balanced

counts, rather than being left to one unrestricted random draw. In this way, the resulting depth dependence can be compared with the optical structure described in Chapter 3 without reducing each depth to one corridor or one special offset from a DOM level.

5.4.3 Zenith

Zenith required a slightly different construction. My earliest tests started from vertical tracks and tilted them by fixed angles, but an inclined or vertical infinite line naturally crosses several ice layers. Unlike the horizontal azimuth and depth scans, the zenith dependence can never be made completely independent of the detector's vertical structure. The goal is instead to keep the spatial sampling comparable as the direction rotates.

For the production scan, I sample azimuth uniformly and place the tracks inside a disk perpendicular to their direction. Sampling uniformly in disk area is important here because a naive uniform radius would not represent the disk area correctly. This construction keeps a controlled transverse region around the selected centre while the zenith angle is varied.

The nominal zenith scan uses 5° bins over the full 0°–180° range. For each event, the azimuth is sampled uniformly and the track support point is drawn from a disk of radius 225 m perpendicular to the sampled direction. The disk is centred at $(0, 0, -200)$ m, a depth chosen to place the reference sample in a comparatively clear region of the ice. This gives favourable photon-propagation conditions for studying the zenith dependence, which is particularly sensitive to the sampling geometry and generally more difficult to reconstruct accurately.

Because the disk rotates together with the track direction, the physical depth distribution of the sampled lines inevitably changes with zenith. To check whether the observed behaviour depends too strongly on this choice of reference position, I produced two additional configurations as robustness controls. The first is detector-centred at $(5.87, -2.52, -23.91)$ m, with a 200 m disk and an independent random seed. The second is centred at $(0, 0, +200)$ m and uses the same 225 m disk, binning, seeds and relative track geometries as the nominal run. Their detailed comparison is presented in the discussion.

Because an up–down difference remained interesting after the production, a more direct paired control was added. For a support point $\mathbf{r}_0$ and direction $\hat{\mathbf{u}}$, the expected response is evaluated again on the identical infinite line but with the opposite orientation $-\hat{\mathbf{u}}$. The two cases share exactly the same geometrical line and differ only in the sense of propagation. This removes much of the ambiguity present when comparing two independently sampled zenith bins.

The control uses 36 azimuths separated by 10°. The paired directions are 0°/180°, 20°/160°, 45°/135° and 65°/115°. For the horizontal case, I also compare a 90° track with the same line at the opposite azimuth, $\phi \to \phi + 180°$. It reuses the perpendicular-disk construction with radius 225 m. Expected charge and expected non-empty DOM multiplicity are first evaluated before Poisson sampling. I then use 100 paired Poisson realisations per line to check that counting fluctuations alone do

not create the asymmetry. A very small reconstruction check was also performed as a technical diagnostic, but its six events per orientation are not used as a statistical reconstruction result.

5.4.4 Detector-centred radial dependence

The final radial production is the detector-wide version of the boundary study developed above. Instead of selecting one physical side, I vary the closest approach of the track to the chosen detector centre and follow the reconstruction from the central region towards and beyond the instrumented volume. The production uses 15 closest-approach radii,

$$R = 0,\ 50,\ 100,\ 150,\ 200,\ 250,\ 300,\ 350,\ 400,\ 450,\ 500,\ 550,\ 600,\ 650,\ 700 \text{ m}. \tag{5.10}$$

For this production, I use the geometrical DOM centroid as the detector centre, $(x_c, y_c, z_c) = (5.87, -2.52, -23.91)$ m. The first run contained ten radii, $R = \{0, 50, 100, 150, 200, 300, 400, 500, 600, 700\}$ m. I later added the five intermediate radii, $R = \{250, 350, 450, 550, 650\}$ m, with the same sampling protocol and 1 000 requested events at each radius. I merged the event-level samples before calculating the summaries, giving a final production of 15 radii and 15 000 requested events. The directions are isotropic and the tangent positions are sampled randomly at each radius. Here, charge, hit multiplicity and success fractions are especially important. If I looked only at the angular error of the successful events, I could miss the more basic fact that the detector is progressively losing information as the tracks move outside the dense instrumented region.

The scan coordinate R and the 400 m response cutoff defined in Section 5.1.1 are different distances. The first is the closest approach of the track to the detector centroid, whereas the second is evaluated separately between that track and every DOM. A radial bin at $R = 400$ m is not a boundary at which the complete detector response is switched off. For the outer bins, however, the cutoff can reduce the set of DOMs allowed to contribute sampled charge. It is consequently part of the simulation definition used to interpret the outer radial behaviour.

The equal number of requested events at each radius is useful for resolving the scan, but it does not represent a population whose closest-approach points are uniform in area. The area weighting was not part of event generation or reconstruction and did not change any simulated event. It was introduced only afterwards, as a post-processing integration of the existing event-level results, to define one reference population from the radial scan. Each simulated radius R_k represents the annulus bounded by the midpoints to its neighbouring radii. If these bounds are denoted by $R_{k,-}$ and $R_{k,+}$, its geometrical weight inside a reference disk of radius R_{ref} is

$$w_k = \frac{R_{k,+}^2 - R_{k,-}^2}{R_{\text{ref}}^2}, \qquad \sum_k w_k = 1. \tag{5.11}$$

The inner boundary is zero, the intermediate boundaries lie halfway between successive simulated radii, and the final boundary is R_{ref}. Since every direction is

already sampled isotropically, these annular weights define a population with isotropic directions and closest-approach points distributed uniformly in the reference disk.

The reconstruction fraction keeps all requested events,

$$f_{\mathrm{reco}}^{\mathrm{ref}} = \sum_k w_k \frac{N_{\mathrm{reco},k}}{N_{\mathrm{gen},k}}, \tag{5.12}$$

whereas the angular distribution is conditional on successful reconstructions. Its mean is

$$\langle \Delta\psi \rangle_{\mathrm{ref}} = \frac{1}{f_{\mathrm{reco}}^{\mathrm{ref}}} \sum_k \frac{w_k}{N_{\mathrm{gen},k}} \sum_{i \in \mathcal{R}_k} \Delta\psi_i, \tag{5.13}$$

where $\mathcal{R}_k$ is the set of reconstructed events at radius R_k. The weighted p_{50}, p_{68} and p_{95} use the same event weights, normalised over the reconstructed population. Their statistical stability is evaluated with 1,000 stratified bootstrap resamples within each radius; failed events remain in the resampling when the reconstruction fraction is recalculated.

The area-weighted directional bias of the full reference population is calculated from the aligned event vectors defined in Section 5.2, using the same annular event weights before their vector mean is formed. It is a coherent displacement of the complete reference population, not an average of the non-negative event-wise angular errors.

The main reference uses $R_{\mathrm{ref}} = 500$ m, an operational acceptance chosen before the clear loss of event survival in the outer part of the scan. It is not interpreted as a spherical IceCube boundary. The same calculation is repeated with $R_{\mathrm{ref}} = 700$ m to show how the reference changes when the complete stress-test range is included. Both use the common $s_E = 1.425$ light-yield convention motivated at 100 TeV, rather than an explicitly propagated and reconstructed 100 TeV muon energy.

Here, one fixed detector reference is shared by the whole boundary production. The local study below uses a different construction in which each selected DOM defines its own centre.

5.4.5 DOM-centred local placement diagnostics

The DOM-centred studies are deliberately kept separate from the detector-wide productions because they answer a different question. The aim here is not to measure a global dependence of IceCube. Instead, tracks are moved around selected DOMs to see how sensitive the reconstruction is to local placement in different parts of the detector.

The primary production used 24 local centres distributed across the detector. This selection is the basis of the two local results presented in Chapter 6: the evolution with transverse distance from a DOM and the comparison between positions above and below it. The selection is not spatially uniform, however. It contains 18 outer, 3 intermediate and 3 central DOMs and is weighted towards the detector periphery.

The unexpectedly large upper tails of this first production motivated a wider 42-DOM diagnostic using the same local construction. It contains 3 central, 21 intermediate and 18 outer DOMs. Its separation into the three detector regions, together with the centre + intermediate subset of 24 DOMs, is used in the Discussion to test how much of the original tail is associated with the population of detector locations. This centre + intermediate subset must not be confused with the primary edge-weighted 24-DOM production: they contain different modules and answer different questions.

The selected modules cover different depths and detector positions. Three representative examples from the exploratory development are shown in Figure 5.14: an outer module, a central module near $z = 0$, and a deeper central module. The deeper example is only a location label; it does not define a separate HQE category.

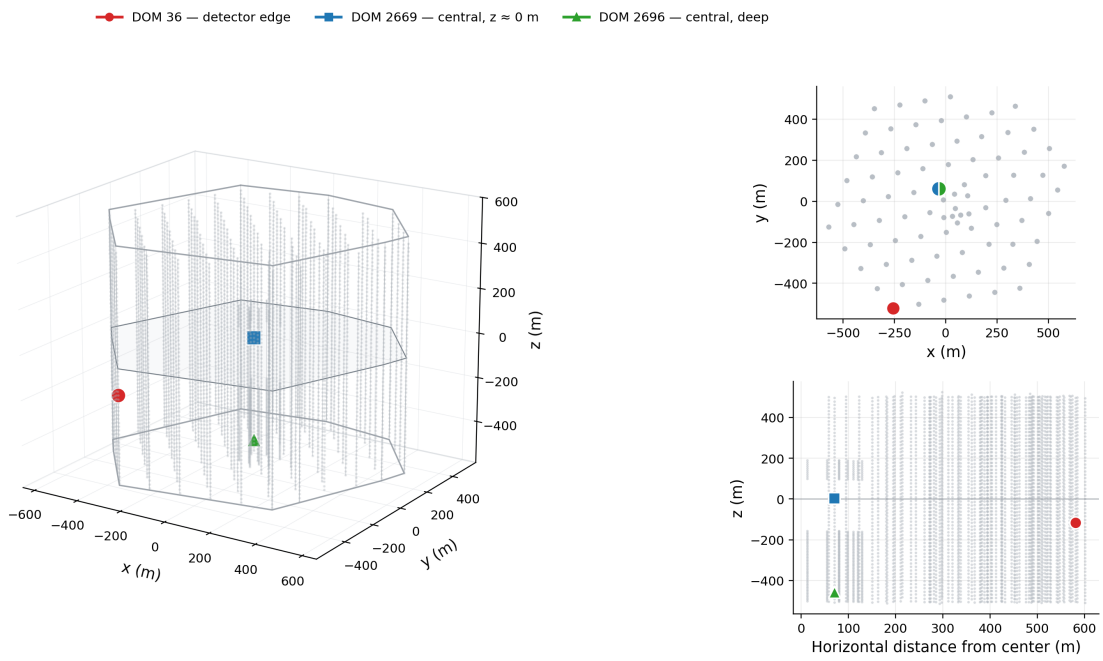


Figure 5.14: Three representative DOM locations from the wider detector sampling, shown in three dimensions, in top view and as a function of horizontal distance and depth. Comparing such positions tests how strongly a local placement diagnostic remains conditioned by the surrounding detector geometry and depth-dependent optical conditions.

Around each selected module, I use two separate constructions. In the horizontal scan, I vary the radial distance from the selected DOM. In the vertical scan, I apply a signed offset above or below the DOM depth. Both use the same local reference scales, but the two coordinates are not varied together. The useful local scale is set by the midpoint between neighbouring modules. Beyond this midpoint, a position described as lying above one DOM would also belong to the below-DOM neighbourhood of the next module, so the two local regions would overlap. The code obtains the relevant neighbour distance directly from the detector-geometry file and uses half of it as the local reference scale.

Both coordinates are expressed through the same dimensionless fractions,

$$f \in \{0.05,\ 0.15,\ 0.30,\ 0.50,\ 0.75,\ 0.90\}, \qquad r = f\, d_{1/2}^{\mathrm{hor}}, \qquad |\Delta z| = f\, d_{1/2}^{\mathrm{vert}}. \quad (5.14)$$

The endpoint is deliberately kept at 90% rather than at the midpoint itself, which preserves an unambiguous association with the selected DOM. For the standard-spacing modules retained in this production, $d_{1/2}^{\mathrm{hor}} = 54.125$ m and $d_{1/2}^{\mathrm{vert}} = 8.5$ m, the latter following from the usual 17 m vertical DOM spacing. Modules close to the denser DeepCore spacing were not selected as local centres because the same numerical scale would no longer describe their immediate neighbourhood.

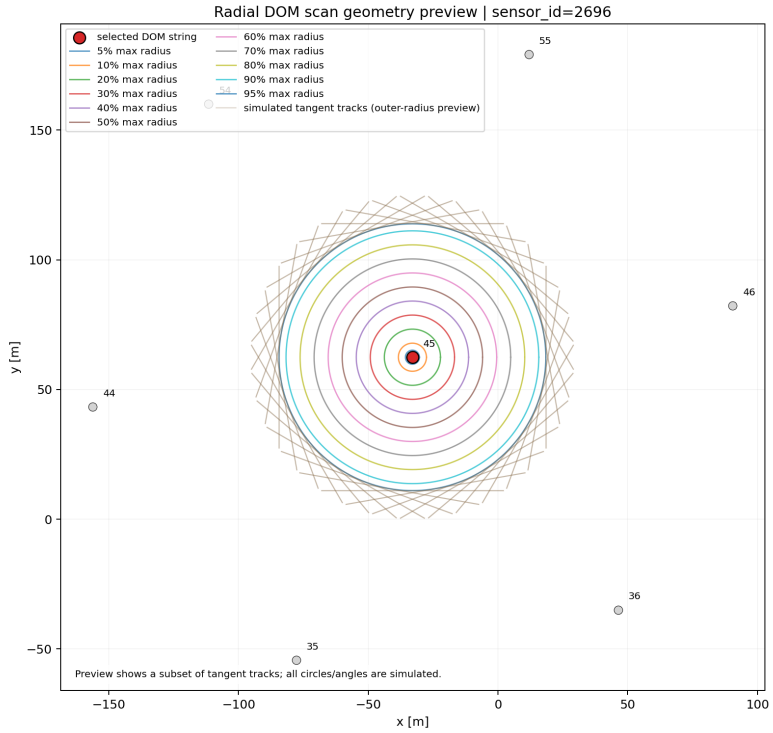
As for the other final scans, a complete radial bin is not represented by one exact position. For each event, I draw the radius uniformly within ± 0.25 m of its nominal value. The polar position is drawn uniformly within one of the 24 adjacent 15° sectors. The percentages in Eq. (5.14) are then used as the common grouping coordinates in the final summaries. Tracks are generated on both sides of the selected module for the vertical scan. Keeping the sign of Δz is important: using only $|\Delta z|$ would average away any possible above–below asymmetry before it could be observed. The two constructions are illustrated in Figure 5.15.

Of course, a "local DOM test" does not mean that the event is reconstructed with one DOM only. The surrounding detector has to remain active: a single hit DOM cannot constrain an infinite track, and removing all the other modules would create a different and largely poorly posed reconstruction problem. The term *local* refers only to the controlled position of the true track relative to one selected DOM.

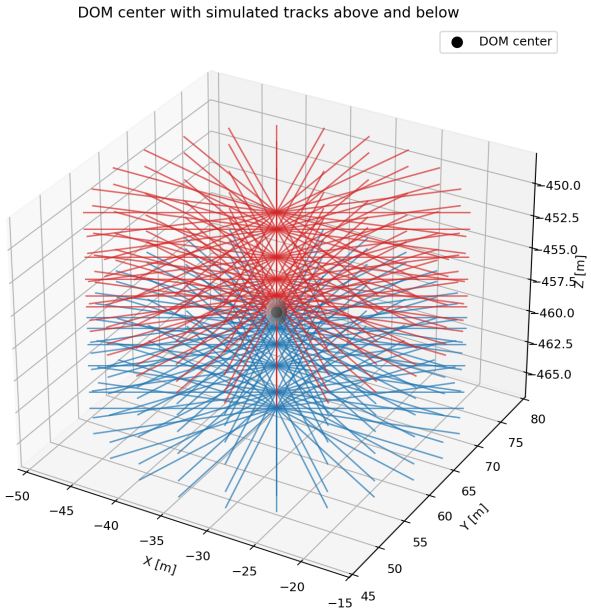
The first versions of this test either mixed horizontal and inclined lines or fixed too many coordinates at once. This made the interpretation unnecessarily difficult: an inclined line crosses several ice layers and can also approach a neighbouring DOM above or below the module chosen as the local centre. A purely horizontal control is easier when the question is specifically the vertical offset.

In the radial protocol, the local radius is the primary scan variable. Each of the six radii contains 24 angular positions separated by 15°, two tangential directions per position and three light realisations per direction. As defined in Section 5.1.2, these are three independent charge and timing draws for the same expected response and line, not three different optical models. This gives 144 requested events per DOM and per radius, or 864 radial events per DOM over the complete scan. In every radial bin, I sample the local vertical offset uniformly in $[-0.25, +0.25]$ m and the true zenith uniformly in $[89°, 91°]$. I keep both distributions unchanged across the six bins, so the local radius remains the main binned coordinate.

For the vertical protocol, every signed offset contains 24 directional orientations and three independently sampled light realisations, giving 72 requested events per DOM and per signed position. At each signed position, I jitter the vertical offset uniformly by ± 0.25 m. I sample the zenith uniformly in $[89°, 91°]$ and the azimuthal axis within its 15° sector before generating the two opposite orientations. The first 24-DOM execution contained the 15, 50 and 90% distances on both sides of each module, or 432 vertical events per DOM. The 5, 30 and 75% distances were added afterwards with the same statistics. The harmonised 24-DOM sample used in the final analysis contains 12 signed vertical positions and 864 vertical events per DOM. The 18 additional DOMs were generated directly with this complete radial and



(a) Top-view annuli and a subset of tangent tracks around DOM 2696.



(b) Three-dimensional track sampling above and below the selected DOM.

Figure 5.15: Geometry of the local DOM-centred diagnostic. The horizontal distance is defined relative to the selected module, while direction, vertical side and offset are varied around it. The surrounding detector remains active in both generation and reconstruction.

vertical protocol.

In both cases, the coordinate of interest is controlled while the nuisance geometry is varied, instead of letting one favourable or unfavourable line represent the whole bin. This brings the local study back to the same rule used for the detector-wide productions.

I first produced the local summaries for only three representative DOMs. I keep these exploratory comparisons because they make the limitation of a single local curve very clear. Figure 5.16 compares the radial behaviour at the detector edge, near the centre and at a deeper central position, while Figure 5.17 shows the same three locations as a function of the vertical offset. The charge, hit multiplicity and reconstruction errors do not evolve in the same way at all three positions. Their differences cannot be attributed to one intrinsic property of the selected DOM; the surrounding detector context still matters.

For the same reason, the signed above–below comparison is kept separately. The DOM 2696 dashboard in Figure 5.18 compares tracks generated above and below the same module under otherwise similar local conditions. Even if the two curves eventually agree, this is still a necessary control: averaging directly over $|\Delta z|$ would remove the sign by construction and would make an asymmetry impossible to detect.

Even with the common randomisation protocol, neither the primary 24-location sample nor the extended 42-location diagnostic is interpreted as a measurement of one isolated DOM property. Each location still combines the local string geometry, depth-dependent ice, containment and, in the lower central region, the greater sensor density produced by the additional DeepCore strings. Since those strings are already part of both the trained model and the simulated geometry, their contribution cannot be switched off and the comparison cannot be interpreted as an HQE measurement. The extended sample and its regional subsets remain discussion controls rather than replacements for the primary local curves.

Table 5.1, placed at the start of the production section, gathers the detector-wide event counts and conditions. The local samples are instead organised jointly by DOM location and local coordinate, so their staged counts are kept separately in Table 5.2.

Stage or analysis population	DOMs	Radial requested	Vertical requested	Total requested
Initial primary execution	24	20,736	10,368	31,104
Vertical complement	24	–	10,368	10,368
Harmonised primary sample	24	20,736	20,736	41,472
Additional production	18	15,552	15,552	31,104
Extended combined sample	42	36,288	36,288	72,576

Table 5.2: Requested event counts for the staged local DOM productions. A complete radial scan contains 144 events per DOM and per radius, while each signed vertical position contains 72 events per DOM; the corresponding complete scans each contain 864 events per DOM. The harmonised 24-DOM and combined 42-DOM rows are the analysis populations obtained after merging the relevant production stages; they are not additional simulations.

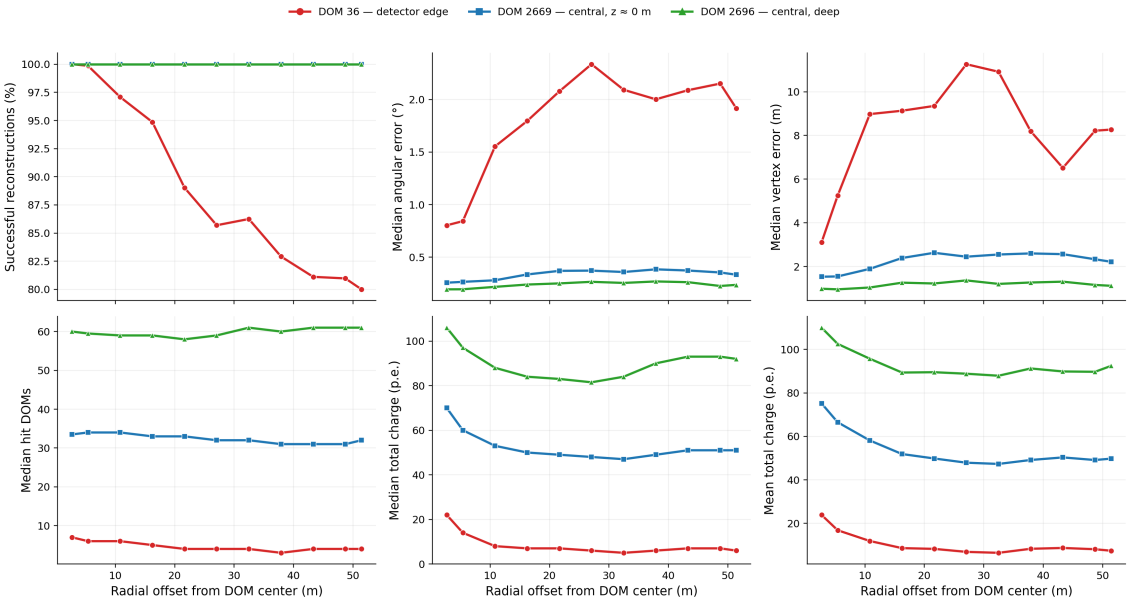


Figure 5.16: Six reconstruction and light observables from the exploratory radial placement test at three representative DOM locations. The common radial coordinate does not make the surrounding string density, containment or ice depth equivalent. The figure motivates the wider final multi-location protocol rather than defining a DeepCore or HQE comparison.

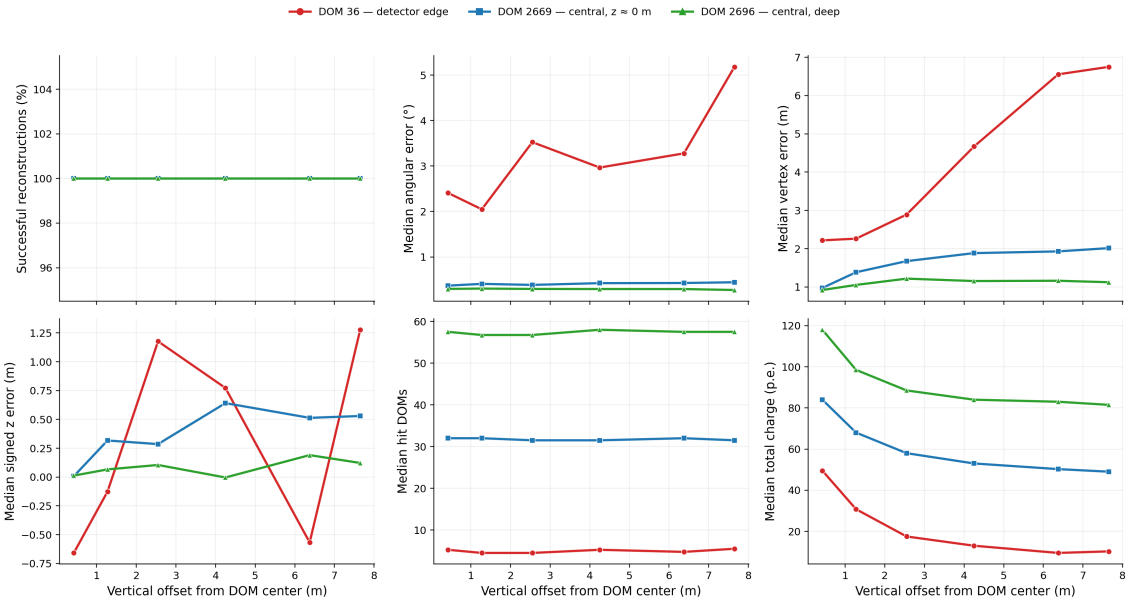


Figure 5.17: Six observables as a function of the magnitude of the vertical offset from the same three representative DOMs. A common vertical scale makes the strong dependence on detector position visible, but does not retain the sign of the displacement.

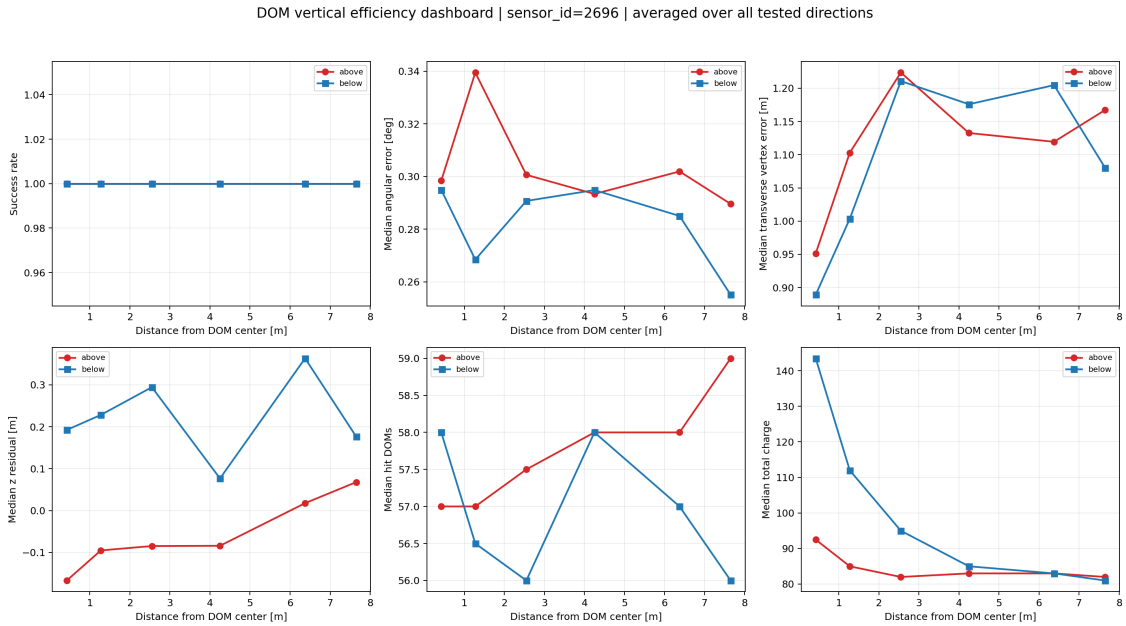


Figure 5.18: Above–below reconstruction diagnostic for DOM 2696, averaged over the tested directions. Success fraction, angular error, transverse spatial error, signed vertical residual, hit multiplicity and charge are kept together so that any vertical difference can be interpreted in terms of both reconstruction and available light. This is a local reconstruction diagnostic, not a measurement of intrinsic DOM efficiency.

Targeted post-processing controls. Because the event-level outputs are preserved, some interpretations can also be tested afterwards without repeating the expensive reconstruction. Three such controls were added after the main productions. First, the extended 42-DOM local sample is separated into its central, intermediate and outer components to determine which detector region contributes most to the upper angular-error tail observed in the local study. The regional uncertainty is estimated with the DOM-level bootstrap described above. Second, the true horizontal tracks of the pooled azimuth production are compared with the 86 string positions through perpendicular track–string distances. Third, for the paired zenith control, the expected charge is aggregated according to DOM depth to check whether the depth distribution carrying the optical information changes with the track direction.

These are post-processing diagnostics of already existing event samples, not additional production scans. They are used in Chapter 7 to test specific interpretations of the local tails, the small azimuth structures near 100° and 280°, and the zenith asymmetry.

5.5 Preservation of event-level results

The reconstruction is the computationally expensive part of the analysis, so I keep the event-level outputs before performing any bin-wise aggregation. This means that once a production is finished, I can change a percentile, inspect a signed residual, compare an error with charge or hit multiplicity, or build a new diagnostic without simulating and reconstructing the same events again. Failed and undetected events remain in the same record as well, which allows the success fractions to be recomputed from the original requested sample.

This separation between event production and later statistical presentation is important for the way the results are treated. The figures are not the primary data of the study; they are different summaries of a preserved event sample. This is also why some of the exploratory plots could be redesigned later without changing the underlying reconstruction. The computational blocks, file organisation, JAX execution and hardware-related adaptations are described in Appendix C; they do not define additional physical intervals or reconstruction observables.

Overall, the methodology is not simply a collection of independent scans. The first versions were useful precisely because they showed where the initial idea of "vary one parameter and keep everything else fixed" breaks down in a discrete detector. A fixed line can end up measuring its own particular position relative to the DOMs more strongly than the variable I intended to study. The final protocols keep the same goal—isolating one dependence at a time—but do so by sampling the nuisance geometry in a controlled and reproducible way. The following chapter applies the common set of angular error, signed residuals, light, hit multiplicity and event-survival quantities to these production samples, with spatial diagnostics added only when they help explain a specific geometrical effect.

Chapter 6

Results

The previous chapter established the method used to test the reconstruction and to separate the effects of the different parameters as far as possible. It also showed that some dependencies cannot be removed completely, since they arise from the geometry and optical structure of IceCube itself. We can now apply this protocol to the final productions and examine the reconstruction level that is reached in each scan. This chapter first presents the main values and the behaviour visible in the figures. The origin of these behaviours, together with the additional tests used to examine them, is discussed in Chapter 7.

For completeness, the full overview dashboards associated with the final production scans are provided in Appendix B. They gather the reconstruction, detected-light and event-survival observables on common axes, while the present chapter retains only the panels needed to describe the main results.

6.1 Reading conventions

Before going through the scans, a few conventions must be recalled. Their full definitions are given in Chapter 5 and Appendix A, but they remain important here because the figures do not all describe exactly the same event population.

The angular error is evaluated event by event for successfully reconstructed events. Instead of reducing each bin to one isolated value, its distribution is described through p_{50}, p_{68} and p_{95}, corresponding to the 50th, 68th and 95th percentiles. The median gives the typical reconstruction level, while p_{68} and especially p_{95} show how far the less favourable events extend. These percentiles must not be read as confidence intervals.

The aligned three-dimensional bias indicates whether the reconstructed directions share a coherent displacement. Its magnitude $b_{3\mathrm{D}}$ is the opening angle between the aligned mean direction and the common true direction; b_θ and b_ϕ give its signed local components, with b_ϕ expressed as an azimuthal arc. Coordinate residuals are kept only as diagnostics. In the local DOM studies, $\langle\Delta\theta\rangle$ and $\langle\Delta\phi\rangle$ are the bin-level means of the zenith and wrapped-azimuth residuals, while b_θ, b_ϕ are reserved for aligned-bias components. Bootstrap resampling estimates the sampling uncertainty.

Angular and spatial errors use successfully reconstructed events, while failed events remain in the denominator of the detection and reconstruction fractions. These two aspects must be read separately, especially in the outer radial bins.

6.2 Depth dependence

The depth scan is considered first because it directly shows how much the reconstruction changes along the vertical optical profile. As shown in Figure 6.1, the most visible feature is the increase in angular error in the interval associated with the dust-rich region. The median changes by more than one degree across the scan, while the upper tail responds even more strongly. The representative extrema are kept in Table 6.1 rather than repeated in the text.

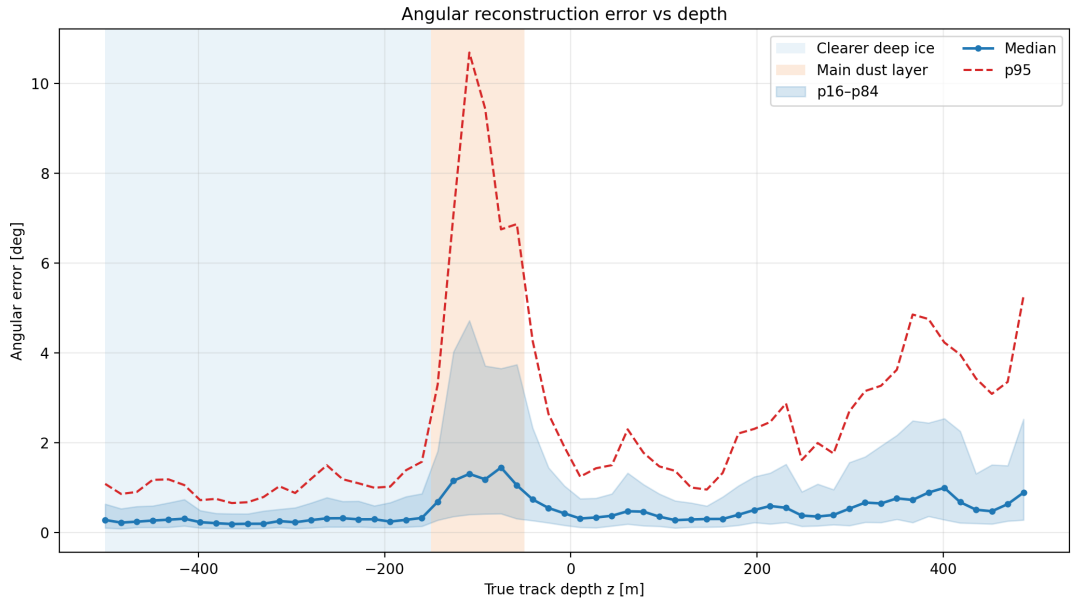


Figure 6.1: Angular-error percentiles as a function of the sampled depth.

Observable	Minimum	Minimum bin	Maximum	Maximum bin
p_{50}	0.19°	−364 m	1.45°	−75 m
p_{68}	0.28°	−364 m	2.31°	−109 m
p_{95}	0.66°	−364 m	10.69°	−109 m

Table 6.1: Angular-error extrema in the final depth scan.

Once this broadening of the angular-error distribution has been established, the signed residuals can be used to check whether it is accompanied by a preferred directional shift. The residuals remain small in comparison with the variation of the positive angular-error metrics, and the bootstrap intervals of their extrema include zero. The main depth dependence is visible in the width and upper tail of the distribution rather than as a resolved one-sided shift. The corresponding bias values and information diagnostics are used in Section 7.3 to examine this interpretation.

The median charge and hit-DOM curves are not repeated here as additional result panels. They are introduced directly in the Discussion, where they are used to test whether the broadening follows a reduction of the information available to the fit.

6.3 Azimuth dependence

The azimuth scan raises a different question, since the measured direction is sampled in optical regions that do not give the same reconstruction level. The global dependence is shown first by pooling the event-level rows of the three final depth-layer productions. This pooling is performed before calculating the percentiles; it is not an average of the three layer summaries. Since the three regions have equal weight, Figure 6.2 must be read as an aggregated azimuth reference rather than as an average weighted by the actual detector event population.

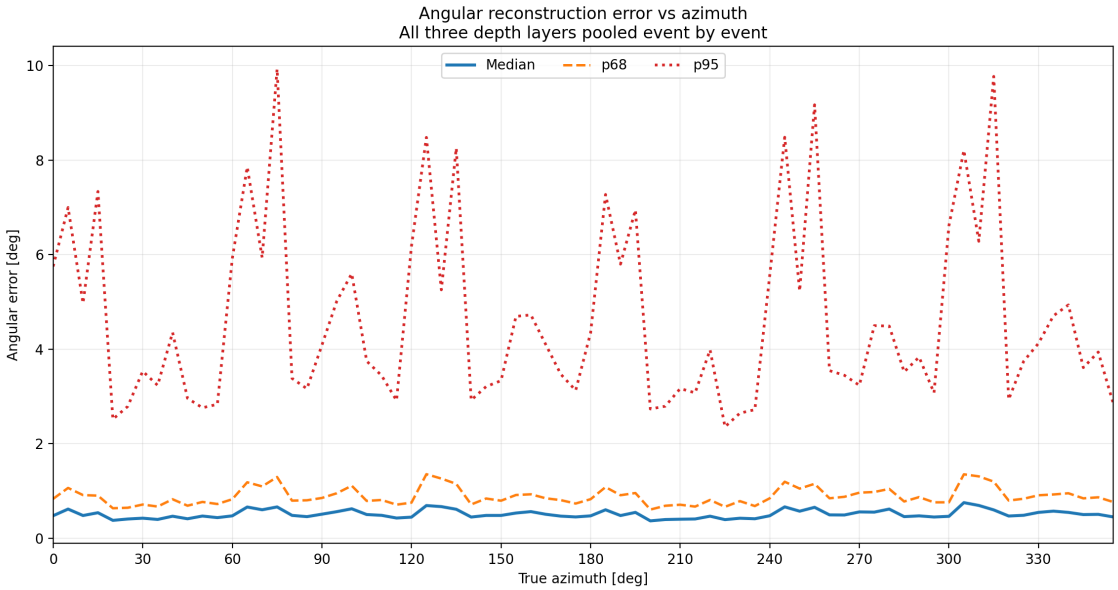


Figure 6.2: Pooled angular-error percentiles as a function of the true azimuth.

The typical reconstruction remains below one degree throughout this pooled scan, while the upper tail varies more strongly. The curve is not flat between its extrema. Repeated higher regions appear near 75°, 125°, 245° and 305°, whereas lower values are found near 20°, 200° and 225°. These positions are stated here as observed structures; their relation to the detector geometry is examined in Section 7.4.

Across the pooled curve, p_{50} spans 0.37° to 0.75°, an absolute change of 0.38°, while p_{95} spans 2.36° to 9.93°, a change of 7.57°. These ranges give the observed scale of the directional modulation. They include extrema close to the principal corridor directions, but do not isolate a corridor-only effect.

The same directional variation is visible in event survival and in the aligned bias. The minimum detection and reconstruction fraction occurs at 130°. In the event-level pooled sample, the largest three-dimensional bias magnitude is 0.66° at 310°, with a 95% bootstrap interval of $[0.34, 1.02]$°. At the same angle, the local zenith

component is $+0.66°$, with $[0.33, 1.01]°$, whereas the local azimuthal-arc component is $+0.05°$, with $[-0.08, 0.21]°$. The displacement at this maximum is carried mainly by the local zenith component. These values, together with the main percentile extrema, are collected in Table 6.2. This keeps the exact values available without interrupting the reading of the curve with a longer list.

Quantity	**Minimum / reference**	**Maximum / largest magnitude**
p_{50}	$0.37°$ at $200°$	$0.75°$ at $305°$
p_{68}	$0.61°$ at $200°$	$1.36°$ at $125°$
p_{95}	$2.36°$ at $225°$	$9.93°$ at $75°$
f_{reco}	91.8% at $130°$	100% in fully reconstructed bins
b_{3D}	–	$0.66°$ at $310°$, CI [0.34, 1.02]°
Local b_θ	–	$+0.66°$ at $310°$, CI [0.33, 1.01]°

Table 6.2: Principal numerical extrema of the pooled azimuth production.

The principal extrema do not describe all the smaller structures visible in Figure 6.2. In particular, the regions around $100°$ and $280°$ require a local string-distance diagnostic before they can be interpreted. That post-processing comparison is kept with the geometrical discussion in Section 7.4 rather than inserted between the main production results.

The pooled curve gives the main global result, but by construction it combines three optical regions. To see what is hidden by this global summary, the same azimuth dependence is also resolved into the three underlying productions. They are centred on nominal depths of $+50$, -100 and -250 m. These values are the centres of the sampled layers, not fixed vertex coordinates, and the depth is sampled uniformly over the corresponding 17 m interval.

This separation changes the reading of the global range. Above the dust and in clear ice, the median remains below $0.71°$ in every azimuth bin. In the dust-layer production, it instead reaches $3.00°$, and the maximum p_{95} rises to $26.77°$. The full ranges are given in Table 6.3; here, the important point is that one pooled percentile can contain event populations with very different reconstruction levels.

Layer	p_{50} **range [deg]**	p_{68} **range [deg]**	p_{95} **range [deg]**
Above dust	0.28–0.70	0.40–1.26	0.97–3.83
Dust layer	0.92–3.00	1.37–5.46	4.59–26.77
Clear ice	0.23–0.45	0.35–0.74	0.82–2.67

Table 6.3: Ranges of the bin-level angular-error percentiles in the three depth-resolved azimuth productions.

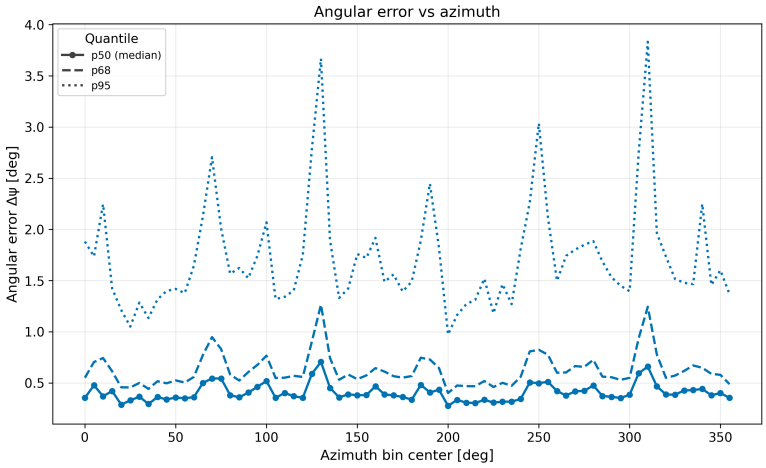
The amount of detected information also differs between the three populations. The median number of hit DOMs ranges from 29 to 35 above the dust and from 46.5 to 59 in clear ice, but only from 3 to 6 in the dust layer. The corresponding charge ranges are 39–50 PE, 67–87 PE and 4–12 PE. All events are detected and reconstructed in every above-dust and clear-ice bin. In the dust layer, both fractions fall to 75.5% at $130°$.

Separating the layers does not remove the local angular structures. Peaks in p_{50} occur near $130°$ and $310°$ both above the dust and in clear ice. In the dust layer,

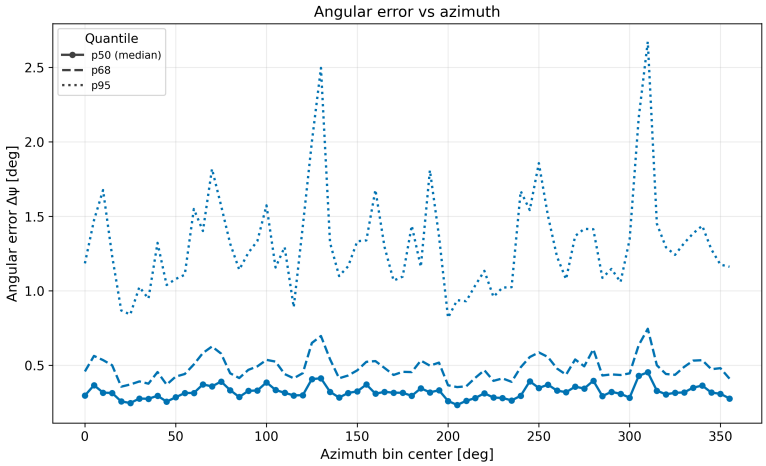
the main maxima are found near 125°, 315°, 75° and 255°. Their repetition across distinct optical regions is visible directly in the production curves; its geometrical interpretation is reserved for Section 7.4.

The aligned directional bias follows the same hierarchy. Its maximum is 0.165° above the dust and 0.101° in clear ice, at 305° and 250°, respectively. In the dust layer it reaches 2.417° at 310°, where the local zenith component dominates the displacement. The complete component curves and their bootstrap intervals are shown in Figure B.7; their interpretation is continued in Section 7.4.

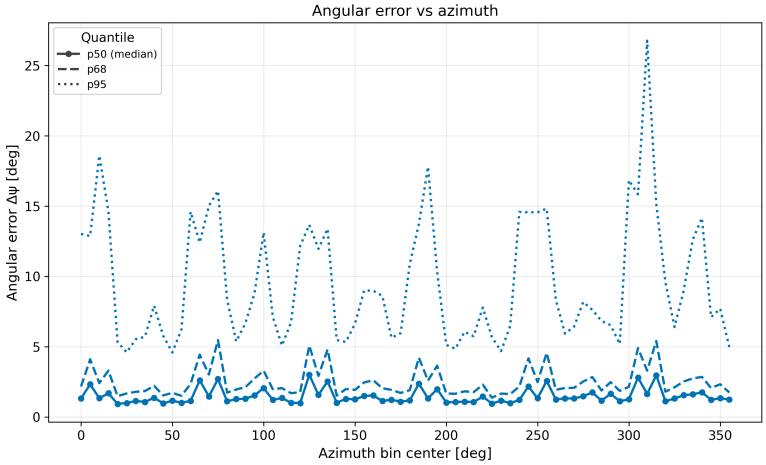
The three curves are finally placed on the same figure to make this difference between global and local behaviour visible. Figure 6.3a shows the production above the dust layer, Figure 6.3b the clear-ice production below it, and Figure 6.3c the dust-layer production. Together they form Figure 6.3.



(a) Above the dust layer.



(b) Clear ice below the dust layer.



(c) Dust layer.

Figure 6.3: Azimuth dependence shown separately for the three depth regions. The panels complement the pooled result by showing how the same directional variable changes with the optical region sampled by the event.

6.4 Zenith dependence

The zenith scan then follows the reconstruction while the track is rotated from one vertical direction to the other. The median angular error ranges from 0.31° at 45° to 1.96° at 175°. The upper tail responds more strongly, with p_{95} increasing from 1.36° at 45° to 14.88° at 180°.

The two sides of the zenith interval do not follow the same evolution. This is not only visible in the angular error, but also in the detected charge, hit-DOM multiplicity and line-to-line distance. Instead of listing every bin, Table 6.4 gives the exact values at five angles that make this progression easier to follow.

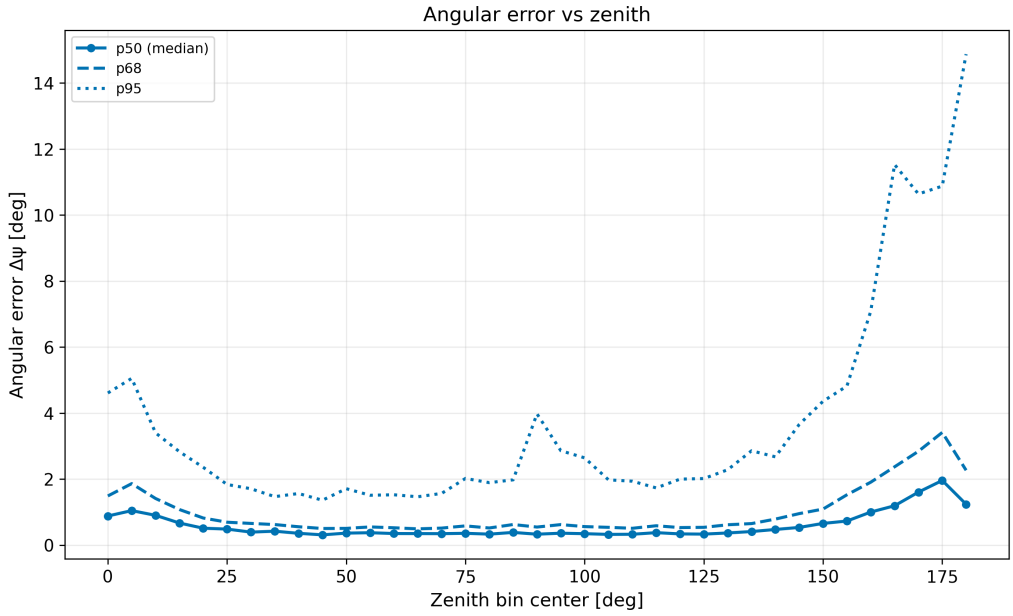
Zenith	p_{50}	p_{68}	p_{95}	Hit DOMs	Charge [PE]	Line distance [m]
0°	0.88	1.49	4.61	44	53	3.64
20°	0.51	0.82	2.36	51	74	1.05
90°	0.33	0.55	3.98	40	63	0.75
160°	1.00	1.91	7.09	20	31	1.65
180°	1.24	2.27	14.88	18.5	24	5.04

Table 6.4: Characteristic bins of the final zenith scan. Angular quantities are in degrees. The line-distance quantity is the median minimum distance between the reconstructed and true infinite lines.

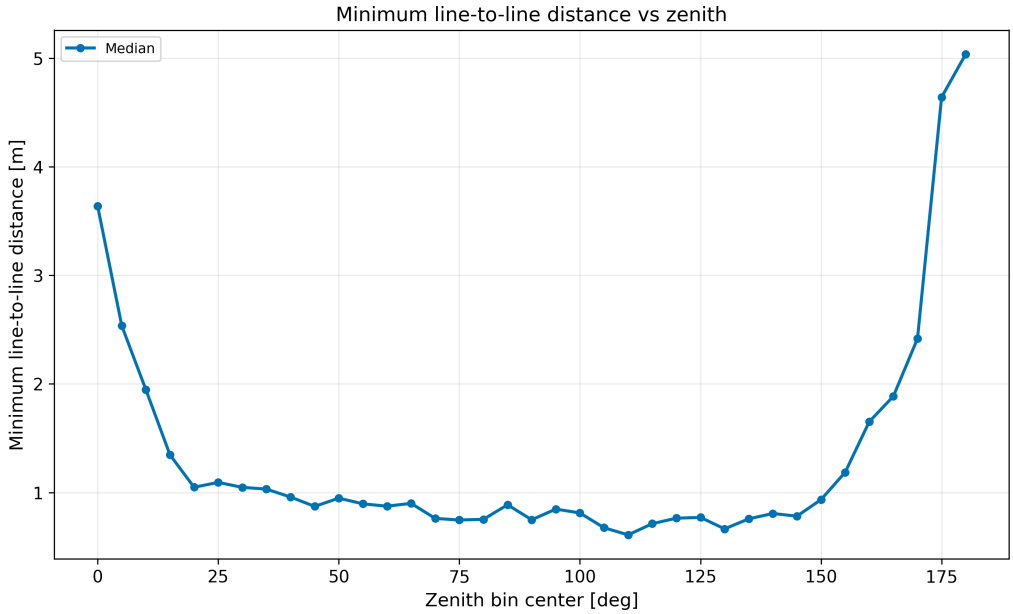
Because this difference appears between opposite sides of the same zenith range, the light response was also evaluated on identical infinite lines with opposite orientations. This control is used in Section 7.5, together with the comparison between the three vertical sampling centres, because its purpose is to test possible explanations of the observed asymmetry rather than to define the primary reconstruction curve.

The spatial reconstruction follows a related but not identical variation. The median line-to-line distance is lowest at 110°, with 0.61 m, and reaches 5.04 m at 180°. This loss of precision does not correspond to a large loss of events: the reconstruction fraction is one at 0°, 20°, 160° and 180°, and remains 99.75% at 90°.

The aligned three-dimensional bias remains below 0.33° in the non-pole bins and reaches 0.329° at 170°, with a bootstrap interval of $[0.072, 0.768]°$. In the pole-centred bin at 180°, its magnitude is 0.673°, with $[0.166, 1.285]°$. The magnitude remains meaningful there, whereas the orientation of its signed local components depends on the chosen roll convention. The complete curve is shown in Figure B.8, and the pole convention is discussed in Section 7.5.1. The angular and spatial results are shown together in Figure 6.4, through Figures 6.4a and 6.4b.



(a) Angular-error percentiles.



(b) Median minimum line-to-line distance.

Figure 6.4: Directional and spatial reconstruction performance in the zenith scan.

6.5 Detector-centred radial dependence

After the two angular scans, the detector-centred radial scan changes the position of the track rather than its direction alone. Its controlled quantity is the minimum distance between the infinite track line and the detector centre. The 15 tested radii cover tracks crossing the central region, tracks approaching the detector edge and tracks passing outside the main instrumented volume. Since IceCube has no unique spherical boundary, this curve must be read as a dependence on distance to the centre, not as a direct distance to one well-defined detector edge.

The radii extend from 0 to 700 m in steps of 50 m. The median error does not increase immediately: its lowest value, 0.41°, occurs at 150 m. It then remains below 1° up to 400 m, first exceeds this value at 450 m and reaches 12.74° at 700 m. The upper tail grows earlier and much more strongly, reaching 59.30° at the largest radius. The value $R = 400$ m here is a distance to the detector centroid; it is not the per-DOM charge-response cutoff described in Section 5.1.1. The two distances should not be identified with one another.

Up to 500 m, this degradation still concerns almost exclusively the quality of the reconstructed events, since the reconstruction fraction remains above 99%. Event survival starts to change clearly farther out: the fraction falls below 95% at 600 m and below 90% at 650 m. The transition from an internal precision loss to a simultaneous loss of events is summarised at representative radii in Table 6.5, together with the detected charge and hit-DOM multiplicity.

R [m]	p_{50}	p_{68}	p_{95}	f_{reco}	Hit DOMs	Charge [PE]
0	0.44	0.69	2.18	100%	26	40
150	0.41	0.68	2.78	99.9%	33	50
450	1.16	2.06	8.30	100%	19	28
500	1.61	3.07	13.60	99.8%	16	25
600	5.40	10.26	37.65	93.0%	8	11
650	9.61	17.64	53.68	77.2%	4	4
700	12.74	23.23	59.30	58.2%	1	1

Table 6.5: Representative bins of the final 15-radius detector-centred production. Angular quantities are in degrees.

The equal event count at each radius is useful for resolving this curve, but it does not correspond to a population distributed uniformly in transverse area. Applying the annular weights of Section 5.4.4 gives the reference values in Table 6.6. For the main acceptance, $R_{\mathrm{ref}} = 500$ m, an isotropic population with closest-approach points uniform in the reference disk has a mean angular error of 1.72°, a median of 0.74° and a reconstruction fraction of 99.97%. The 95% bootstrap intervals are $[1.65, 1.79]^\circ$ for the mean and $[99.93, 99.99]\%$ for the reconstruction fraction.

Extending the acceptance to the complete 700 m stress-test range changes the population itself. The mean then becomes 5.76°, with a 95% bootstrap interval of $[5.54, 5.96]^\circ$, while the reconstruction fraction becomes 93.03%, with $[92.58, 93.48]\%$. The corresponding median is still only 1.47°, but $p_{95} = 27.46^\circ$. The difference between the mean and median reflects the broad outer tail. These two rows are not competing estimates of one universal resolution; they describe two explicitly different radial acceptances.

R_{ref} [m]	Mean $\Delta\psi$	p_{50}	p_{68}	p_{95}	b_{3D}	f_{reco}
500	1.72	0.74	1.32	6.32	0.13	99.97%
700	5.76	1.47	3.55	27.46	0.49	93.03%

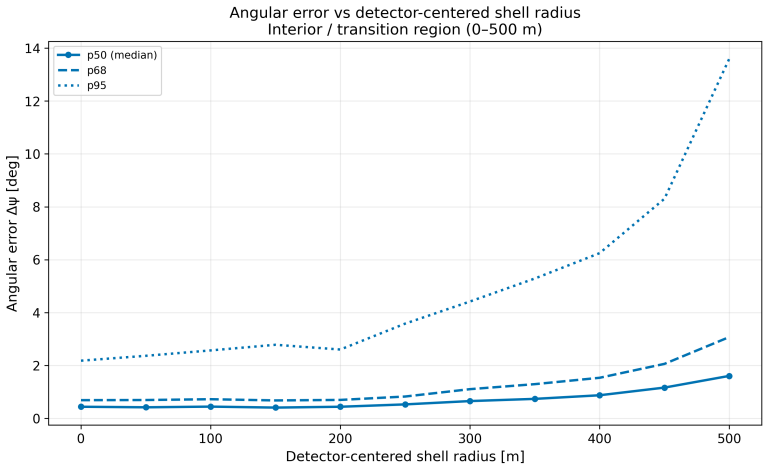
Table 6.6: Area-weighted references obtained from the isotropic detector-centred radial production. Closest-approach points are uniform in a disk of radius R_{ref}. Angular quantities are in degrees and are conditional on successful reconstructions; f_{reco} retains all requested events. The light-yield convention is the common $s_E = 1.425$ setting motivated at 100 TeV.

For the 500 m reference, the area-weighted 3D bias of the reference population is 0.13°, with a bootstrap standard deviation of 0.03° and a 95% bootstrap interval of $[0.07, 0.19]^\circ$. It becomes 0.49°, with $[0.31, 0.69]^\circ$, over the complete 700 m range. These values describe the coherent displacement after the population has been weighted and remain much smaller than the corresponding distribution widths. They should not be confused with the bin-wise outer result reported next.

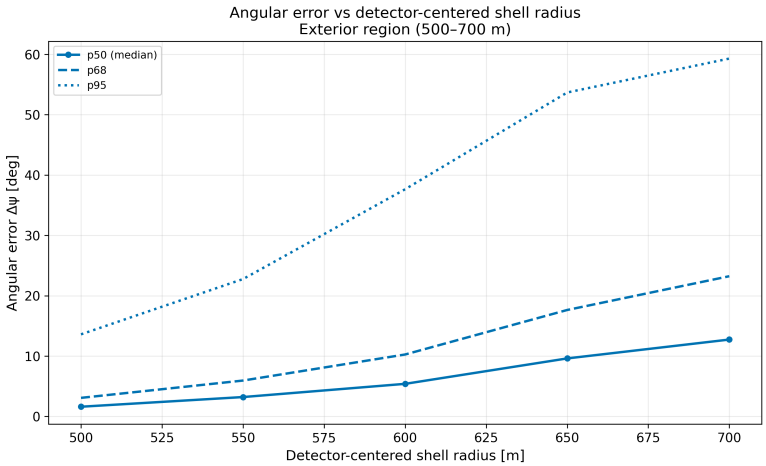
The aligned directional bias remains secondary over most of the radial range and becomes visible mainly in the outermost bins. At 700 m, $b_{\mathrm{3D}} = 3.13^\circ$, with a 95% bootstrap interval of $[1.56, 4.86]^\circ$. The displacement is dominated by the local zenith component, $b_\theta = +3.13^\circ$ with $[1.45, 4.81]^\circ$, while the local azimuthal-arc component is -0.19° with $[-1.49, 1.16]^\circ$. The outer degradation combines a coherent local zenith shift with a much broader distribution of reconstructed directions. The full radial bias curve is shown in Figure B.10.

Because the large outer errors would compress the internal part of one global curve, the two ranges are shown directly through the interior and exterior views in Figures 6.5a and 6.5b, grouped in Figure 6.5. The event fractions are then reported in Figure 6.6, where the loss of event survival can be distinguished from the preceding loss of precision.

Only two distinct traces are visible in Figure 6.6: $N_{\mathrm{det}} = N_{\mathrm{reco}}$ at every radius, and consequently $f_{\mathrm{det}} = f_{\mathrm{reco}}$. The outer event loss comes from realisations with no detected hit, rather than from additional fit failures among the detected events.



(a) Interior range.



(b) Outer range.

Figure 6.5: Zooms on the internal and external parts of the detector-centred radial scan.

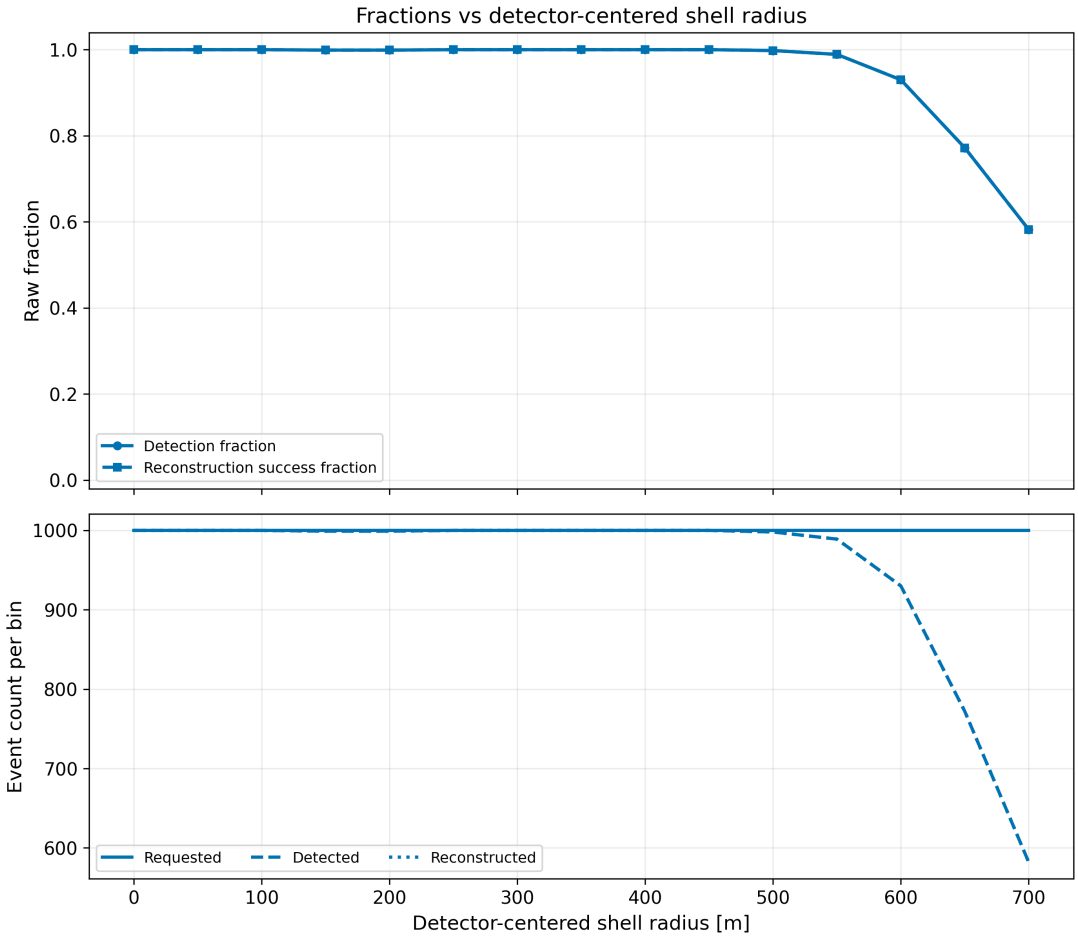


Figure 6.6: Detection and reconstruction fractions and the corresponding event counts as a function of distance to the detector centre. The detected and reconstructed populations are identical in every radial bin, so each pair of curves is exactly superimposed.

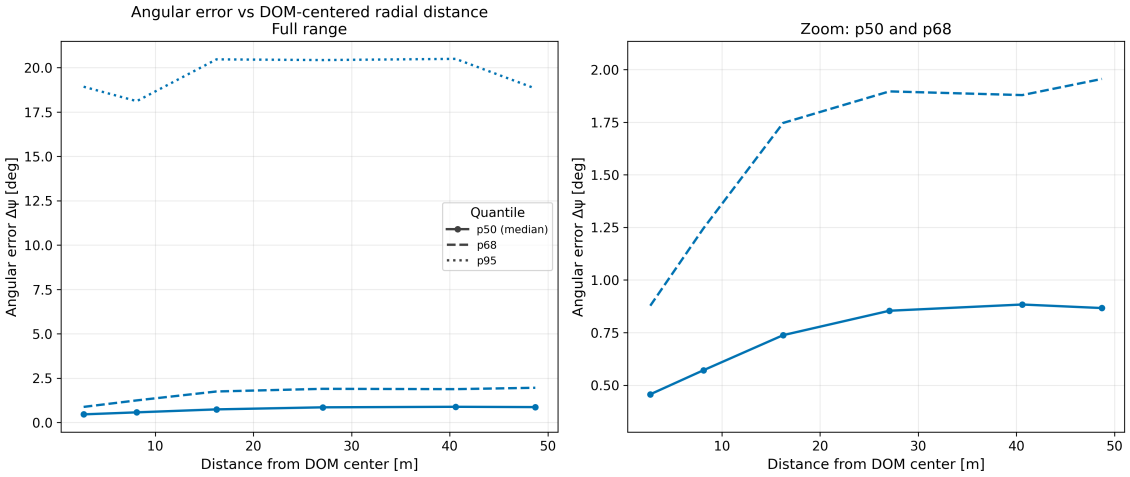
6.6 Local DOM-centred geometry

The detector-wide scans provide the main reconstruction result, but they cannot show how strongly a controlled placement relative to the nearest modules can change it. The primary local production averages over 24 selected DOMs and is organised around two questions. First, how does the reconstruction evolve when the track is moved transversely away from the selected DOM? Second, does placing the track above or below the same DOM produce a visible asymmetry? The six nominal coordinates in both scans correspond to 5, 15, 30, 50, 75 and 90% of the relevant horizontal or vertical mid-distance defined in the methodology.

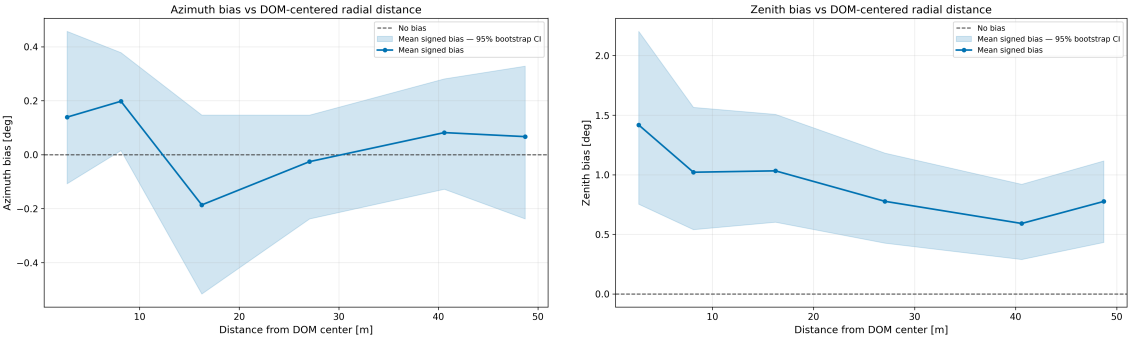
This is still an average over several detector locations rather than a result for one favourable module. Its composition is weighted towards the detector edge, with 18 outer, 3 intermediate and 3 central DOMs. That dependence is not resolved further here: the extended 42-DOM production and its regional split are used in Section 7.7 to understand the large upper tails visible in the primary curves.

6.6.1 Distance from the selected DOM

Figure 6.7 follows the angular error and the two signed coordinate residuals as the transverse distance increases. Across the six radial coordinates, p_{50} lies between 0.46° and 0.88°, p_{68} between 0.88° and 1.96°, and p_{95} between 18.11° and 20.49°. The reconstruction fraction varies from 97.83% to 100%. Over the same coordinates, the median detected charge is between 27 and 40 PE and the median hit-DOM multiplicity between 17 and 19. The typical error remains much smaller than the upper tail, while the mean and maximum values of p_{95} remain close to twenty degrees over the tested range. The mean zenith residual is consistently more visible than the mean azimuth residual, which remains closer to zero and changes sign across the tested positions. This is the main observation that motivates the later population check; it is not yet an explanation of the tail.



(a) Angular-error percentiles.



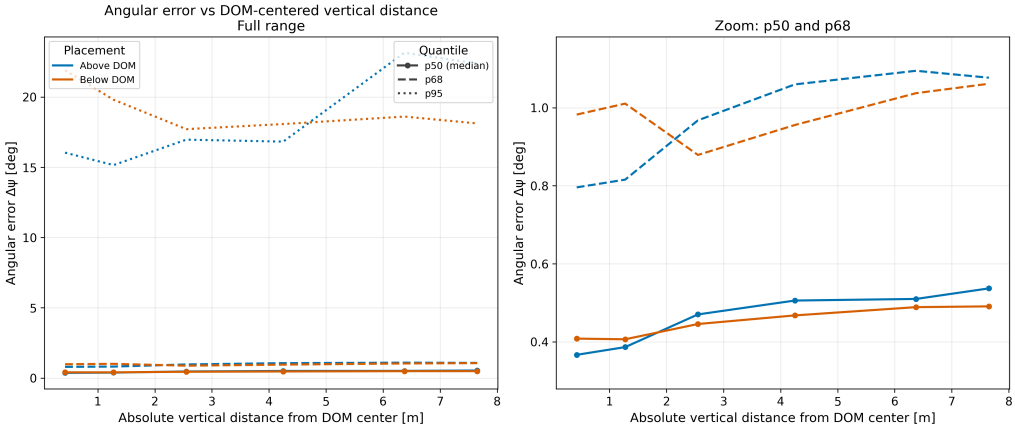
(b) Mean signed azimuth coordinate residual.

(c) Mean signed zenith coordinate residual.

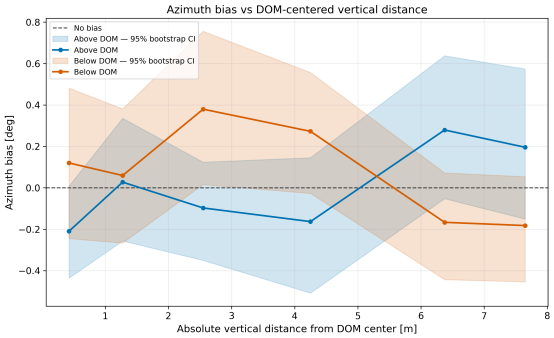
Figure 6.7: DOM-centred radial result for the primary 24-DOM production.

6.6.2 Above–below placement

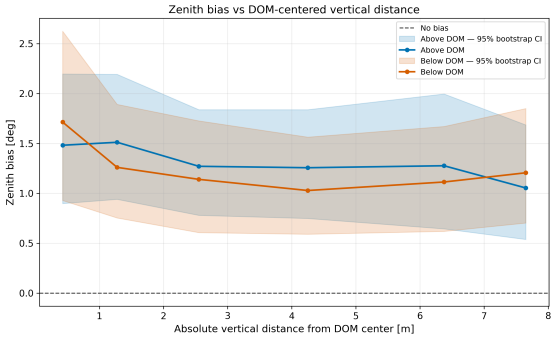
The second scan retains the sign of the vertical displacement instead of combining the two sides through $|\Delta z|$. The curves in Figure 6.8 show that the two placements do not produce a uniform one-sided improvement across every observable. At the two smallest offsets the reconstruction is marginally better above the selected DOM, after which the ordering reverses and the below-DOM placement becomes slightly better. At 0.425 m, the above-DOM sample gives $p_{50} = 0.367°$, $p_{68} = 0.796°$ and $p_{95} = 16.04°$, whereas the values below the DOM are 0.408°, 0.982° and 21.90°. The median charge is 46 PE above and 99 PE below the DOM, while the median hit-DOM multiplicity is 19 and 18, respectively. The same ordering remains at 1.275 m: the median charge is 39 PE above and 65 PE below, while p_{95} is 15.15° and 19.81°, respectively. Their differences remain modest for the typical events but become more visible in the upper tail and in the coordinate residuals. The mean zenith residual stays positive for both placements and is slightly larger above. This common sign is not an above–below reversal: translating the support point vertically does not reverse the track direction, and both samples use the same near-horizontal zenith distribution. The azimuth point estimates have opposite signs with nearly mirrored magnitudes, but every corresponding 95% bootstrap interval includes zero. The mean azimuth residual is consistent with zero in this scan. The quantitative interpretation of the mean zenith residual must retain the detector population used to build the average.



(a) Angular-error percentiles above and below the selected DOMs.



(b) Mean signed azimuth coordinate residual.



(c) Mean signed zenith coordinate residual.

Figure 6.8: Vertical DOM-centred result for the same primary 24-DOM production.

Across the complete vertical scan, p_{50} lies between 0.37° and 0.54°, p_{68} between 0.80° and 1.09°, and p_{95} between 15.15° and 23.15°. Every requested event is detected and reconstructed. Table 6.7 summarises the upper tails and largest absolute mean coordinate residuals of the two primary local scans.

Scan	**Mean** p_{95}	**Maximum** p_{95}	**Maximum** $\lvert\langle\Delta\theta\rangle\rvert$	**Maximum** $\lvert\langle\Delta\phi\rangle\rvert$
Radial	19.54	20.49	1.42	0.20
Vertical	18.73	23.15	1.71	0.38

Table 6.7: Upper-tail and signed-coordinate-residual summary of the primary 24-DOM local production. All angular quantities are in degrees. The last two columns report the largest absolute bin-level means of $\Delta\theta$ and wrapped $\Delta\phi$; their bootstrap intervals are shown in the corresponding figures.

6.7 Cross-scan synthesis

The radial integration provides one angular reference once the population and its geometrical weighting are fixed. The other scans show why this value cannot be transferred unchanged to every detector configuration. The depth scan first establishes the variation across the vertical optical profile. The azimuth scan then shows that a global directional curve retains local structure, while its separation by depth reveals how different the event populations inside that curve can be. The zenith scan adds a difference between opposite sides of the directional range, even though nearly all events are still reconstructed. Moving the track away from the detector centre finally produces a transition from a loss of precision to a loss of event survival.

The local DOM scans complete this picture from another scale. Their purpose is not to replace the global result with a catalogue of local cases. They show instead that a detector-wide average cannot be treated as a performance that would remain unchanged at every position in IceCube. In the primary local sample, this appears mainly through upper tails that are much larger than the typical error. Chapter 7 tests whether that difference is associated with the composition of detector regions, while keeping the two original local questions separate from the later diagnostic. This provides the basis for discussing the optical, geometrical and containment contributions without repeating a second table of values already given in the individual sections.

Chapter 7

Discussion

In this chapter, I return to the results presented in Chapter 6 in order to understand what sets the reconstruction level in each scan. The objective is not to repeat every numerical extremum, but to connect the observed changes to the amount of optical information, the detector geometry and the event population on which each curve is built.

The discussion follows the same progression as the methodology. I first clarify what the theoretical limit means in the matched-model setup and why a global result can still depend on local detector conditions. I then consider the four main scans before returning to the DOM-centred diagnostics, which help explain the large tails seen in the initial local production. The final sections bring these observations together and state the limits of the interpretation.

7.1 Meaning of the theoretical reconstruction limit

Before interpreting the individual scans, the first question is how well an infinite-muon reconstruction can perform under perfect modelling of the continuous emission within the test. As defined in Sections 5.1 and 5.1.3, generation and reconstruction use the same detector geometry, optical-medium description, expected-charge and timing responses, photon-detection calibration and infinite-muon hypothesis. No mismatch is deliberately introduced between the two stages. This matched-model closure test asks how accurately the reconstruction can recover the track that generated the event under this common set of assumptions.

This favourable setup does not imply a zero reconstruction error. The fit does not receive the continuous light field predicted by the model, but a finite and stochastic set of DOM observations. The resulting constraint still depends on the amount of detected light, the number and spatial distribution of hit DOMs, the discrete detector geometry and the numerical properties of the optimisation. In this sense, the "theoretical limit" is conditional on the tested event population and on the way it is sampled. A single reference value can still be reported, provided that this population and its weighting are defined. The area-weighted radial result in Table 6.6 is one such reference. It is not an intrinsic constant of TriplePandel that can be applied to another event selection without modification.

The results are better understood as several reconstruction cases. Favourable, well-illuminated interior configurations give a sub-degree median error, whereas the dust-rich region, the outer detector-centred scan and some local outer-DOM populations develop much larger tails or reduced event survival. They belong to the same theoretical-limit study, but they do not provide the fit with the same information.

The same distinction must be kept when comparing these results with existing IceCube reconstruction studies. Published performances are obtained for other event populations and include effects that are deliberately absent from the present closure test. In particular, conventional high-energy reconstructions are sensitive to the non-uniform stochastic energy-loss pattern of real muons, whereas the model studied here uses the infinite-muon / uniform-loss approximation [5, 6, 19]. The sub-degree values found in favourable configurations are matched-model benchmarks and should not be presented as a replacement for an experimental IceCube angular-resolution curve, but as a limit on what can be achieved using only the continuous light emission of the infinite-muon model.

This definition also explains why the global and local scans have to be read together. A global scan provides the response for a declared sampling distribution, but the averaging can hide detector-dependent effects. The azimuth results separated by depth region and the DOM-centred studies do not introduce competing definitions of the performance; they show which local dependencies enter the broader result.

7.2 Global averages and local detector dependence

With this definition in mind, the global scans remain the principal answer of the work. They average over a declared distribution of nuisance variables and are less dependent on one favourable track. This does not make the result independent of position inside IceCube. These azimuth and DOM-centred controls recover the local dependencies hidden inside the broader curves, without replacing the global result by a catalogue of individual configurations.

The event counts, fixed coordinates and sampling rules for the detector-wide productions are collected in Table 5.1. The discussion below refers to those declared populations rather than to an unspecified detector average.

The nuisance variables were randomised for the same reason. Randomisation does not make IceCube continuous, nor does it remove the physical dependence on its geometry. It reduces the risk that one accidental local arrangement is taken as the general behaviour of the model. Once the production curve has been built, resolving it by depth or detector region then shows which effects were averaged together.

Azimuth gives a direct example of this distinction. Translating a track does not change the definition of its azimuth, which remains the angle of the direction vector and is independent of the detector centre. It changes the position of that direction relative to the neighbouring strings and DOMs. An azimuth scan based on one fixed line is consequently a local test, whereas the production scan estimates the response after averaging over the declared transverse-position distribution and depth layer.

Across all scans, three detector-dependent ingredients make any global number

dependent on the declared population: the depth-dependent optical medium, the local DOM/string geometry and the loss of containment as the track moves away from the detector centre. They are not isolated in the same way by every protocol. Depth is already the controlled variable of the depth scan, while the zenith and detector-centred radial scans intentionally average over the other sampled coordinates. Splitting these productions again would change the population being studied rather than provide the same direct control as the azimuth layer split or the DOM-selection comparison.

The primary 24-DOM production makes this last point particularly clear because its composition is weighted towards the detector edge. The wider 42-DOM test was added to determine whether the large upper tails remain with a broader set of locations and, if so, which detector regions produce them. Since this is a diagnostic of the main results, it is discussed after the four production scans in Section 7.7.

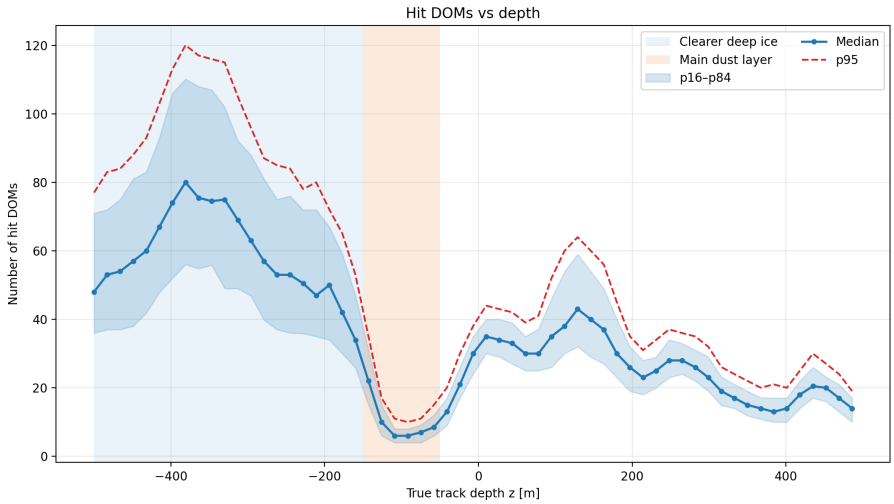
7.3 Depth dependence and available optical information

The depth protocol uses 59 horizontal-track layers from $z = -500$ to $+486$ m, with 200 events per layer, balanced azimuth strata over the full circle and randomised positions inside a 200 m lateral radius (Section 5.4.2). Its angular errors are shown in Figure 6.1 and the representative extrema in Table 6.1. This is the most direct scan because depth also determines the optical environment seen by the track. The best depth bin has $p_{50} = 0.19°$ and $p_{95} = 0.66°$ at $z = -364$ m, whereas the largest median reaches $1.45°$ at $z = -75$ m and the largest p_{95} reaches $10.69°$ at $z = -109$ m. These maxima lie in the interval associated with the dust-rich ice.

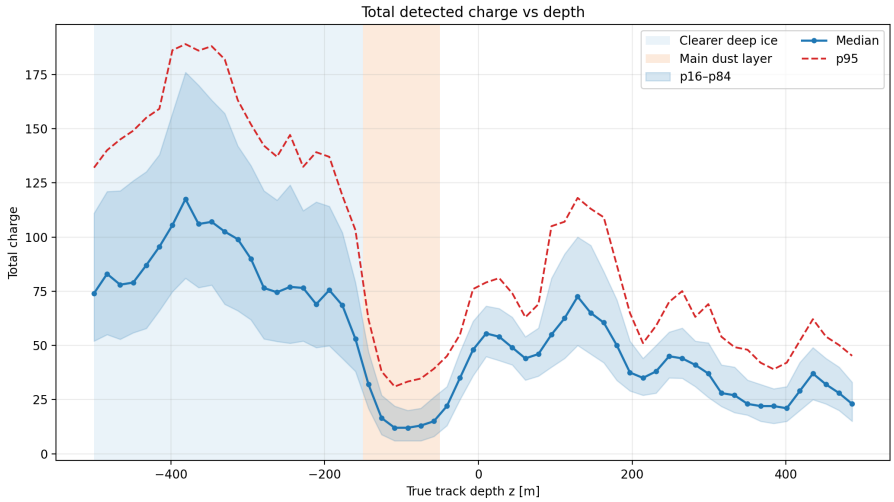
This correspondence is consistent with independent measurements of the South Pole ice. Absorption and scattering vary strongly with depth, and the 2000–2100 m dust layer is less transparent than the clearer ice below it; DeepCore was deliberately placed largely in this deeper, clearer region [22, 36]. However, the scan is not an independent measurement of these ice properties. Their depth dependence is already encoded in both generation and reconstruction. What can be tested here is how the optical variation contained in the common model changes the information available to the fit.

Figure 7.1 combines the hit-DOM count (Figure 7.1a) and total detected charge (Figure 7.1b). These curves connect this optical variation to the reconstruction. The region with the poorest angular performance is also the one in which less information reaches the fit. The relation is not simply proportional, since the likelihood also uses the spatial distribution and timing of the hits. The result supports a degradation associated with reduced optical information, but it does not define a universal charge threshold below which reconstruction suddenly fails.

The signed residuals complete this interpretation. Even at their extrema, the event-level bootstrap intervals include zero. Depth mainly broadens the angular-error distribution in this production rather than producing a resolved one-sided shift of the reconstructed direction.



(a) Median number of hit DOMs.



(b) Median total detected charge.

Figure 7.1: Detected-information diagnostics used to interpret the depth dependence of the angular-error distribution.

7.4 Azimuthal structure and detector geometry

The azimuth production uses horizontal tracks ($\theta = 90°$), 72 bins of 5°, 400 events per bin and per depth, a common ±200 m transverse band, and 17 m layers centred at $z = +50, -100$ and -250 m (Section 5.4.1). The pooled angular errors and their principal extrema are reported in Figure 6.2 and Table 6.2; the comparison by depth region is reported in Figure 6.3. Azimuth is less direct to interpret because the same directional variable was tested in three optical regions with very different reconstruction levels. The pooled curve first shows which structures remain after the three productions are combined at event level. The separate depth-region curves then show how strongly their magnitude depends on the optical region. The first is a broader reference than any individual layer, but it is not weighted by the natural detector event population. These two views answer successive questions rather than providing interchangeable summaries.

With this distinction kept, the first possible explanation to examine is the corridor geometry identified during the exploratory scans. Neighbouring standard-string pairs define three principal axial families at approximately 9.32°, 68.79° and 129.19°, separated by about 60°. The prominent 130°/310° structure differs by only 0.81° from the 129.19° family modulo 180°, and the 75°/255° structure is about 6.21° from the 68.79° family. The alignment supports a geometrical contribution to these features, although it cannot establish that the corridors are their only cause. Figure 7.2 shows the detector layout and the tested axes. The 100° axis is added only as a diagnostic direction and is not treated as a fourth corridor family.

The size of the directional variation can be read from the pooled curve before assigning it to a particular cause. Between its extrema, p_{50} changes by 0.38°, from 0.37° to 0.75°, while p_{95} changes by 7.57°, from 2.36° to 9.93°. The upper values occur in structures compatible with the principal corridor directions. These spans quantify the observed modulation of the reconstruction, but they remain an upper scale for the corridor contribution rather than a causal measurement of that contribution alone.

The smaller structures around 100° and 280° do not follow the same corridor correspondence and were tested separately. The event-level post-processing used for this comparison is summarised in Table 7.1.

For each event in this diagnostic, the true horizontal line is projected into the xy-plane and its perpendicular distance to each of the 86 string coordinates is calculated. "Nearest string" is the smallest of those distances. "Median of 3 nearest" first takes the median of the three smallest distances for one event—equivalently, the second-nearest-string distance—and then reports the median of that event-level metric within the azimuth bin. Each row aggregates the events in one pooled 5° bin across the three depth productions; it is not one simulation. Because the true azimuth and transverse position are sampled event by event, the distance to a string is not held constant within a row.

At 100°, the median distance to the nearest string reaches 4.30 m, compared with 3.32 and 2.88 m in the neighbouring 90° and 110° bins. The opposite side follows the same pattern: the distance is 4.31 m at 280°, compared with 3.41 and 2.88 m

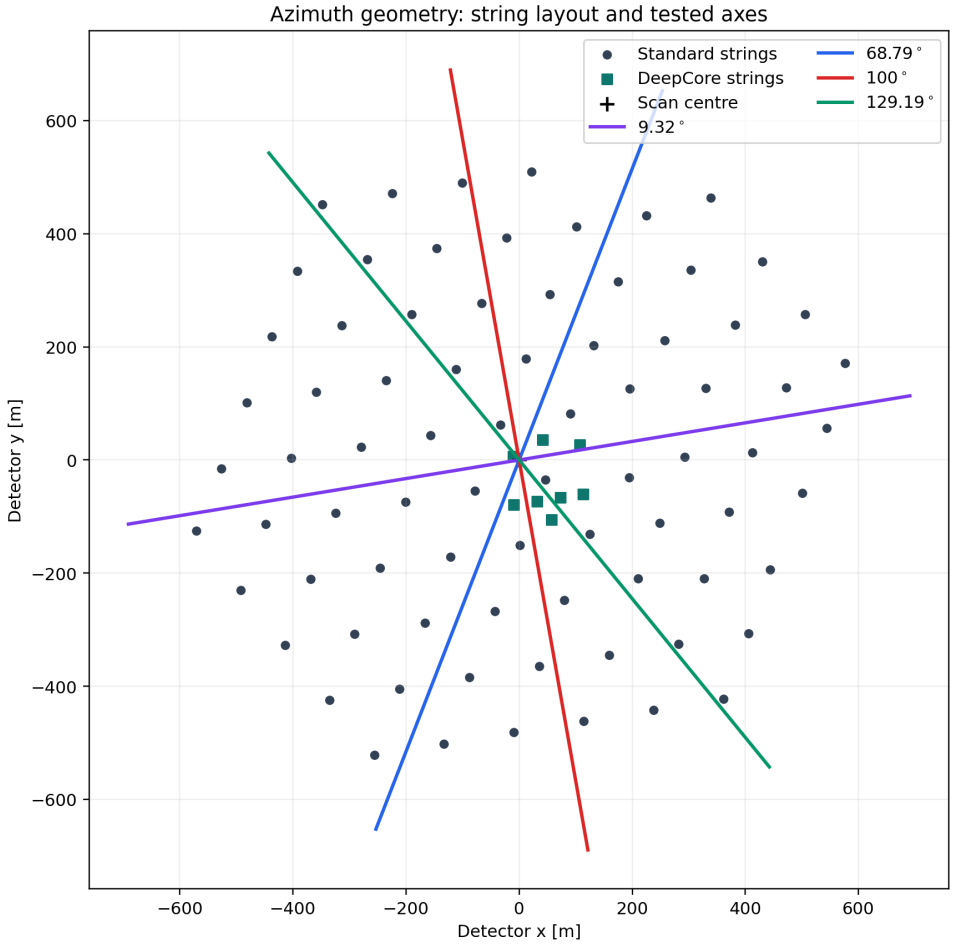


Figure 7.2: Top view of the IceCube string geometry with the three principal axial families used in the corridor comparison and the additional 100° diagnostic axis. Opposite directions are equivalent modulo 180° for this geometrical comparison.

Azimuth	Nearest string [m]	Median of 3 nearest [m]	Hit DOMs	Charge [PE]	p_{50}	p_{95}
90°	3.32	7.84	28.5	43	0.51	4.08
100°	4.30	9.72	28.0	41	0.62	5.60
110°	2.88	7.68	30.0	45	0.49	3.44
270°	3.41	8.15	29.0	42	0.56	3.24
280°	4.31	9.68	29.0	43	0.62	4.49
290°	2.88	7.96	30.0	46	0.47	3.83

Table 7.1: String-distance diagnostic for the smaller pooled-azimuth structures. "Nearest string" is the event-level minimum perpendicular distance to an IceCube string, while "Median of 3 nearest" uses the median of the three smallest distances. Each row is a bin-level median over the pooled event sample. Angular-error quantities are in degrees. The comparison uses the existing pooled production rather than a new reconstruction run.

at 270° and 290°. The angular-error distribution broadens at the same positions. Relative to the average of the neighbouring bins, p_{50} and p_{95} increase by 25.4% and 48.8% at 100°, and by 19.7% and 27.0% at 280°. By contrast, the broader count of strings within 100 m changes much less. These two structures are more naturally interpreted as local consequences of the discrete string geometry than as a newly identified corridor family. Figure 7.3 extends the comparison to the full azimuth range. It shows in particular that the nearest-string metrics contain sharper local variations than the counts within fixed radii, in agreement with the event-level values above.

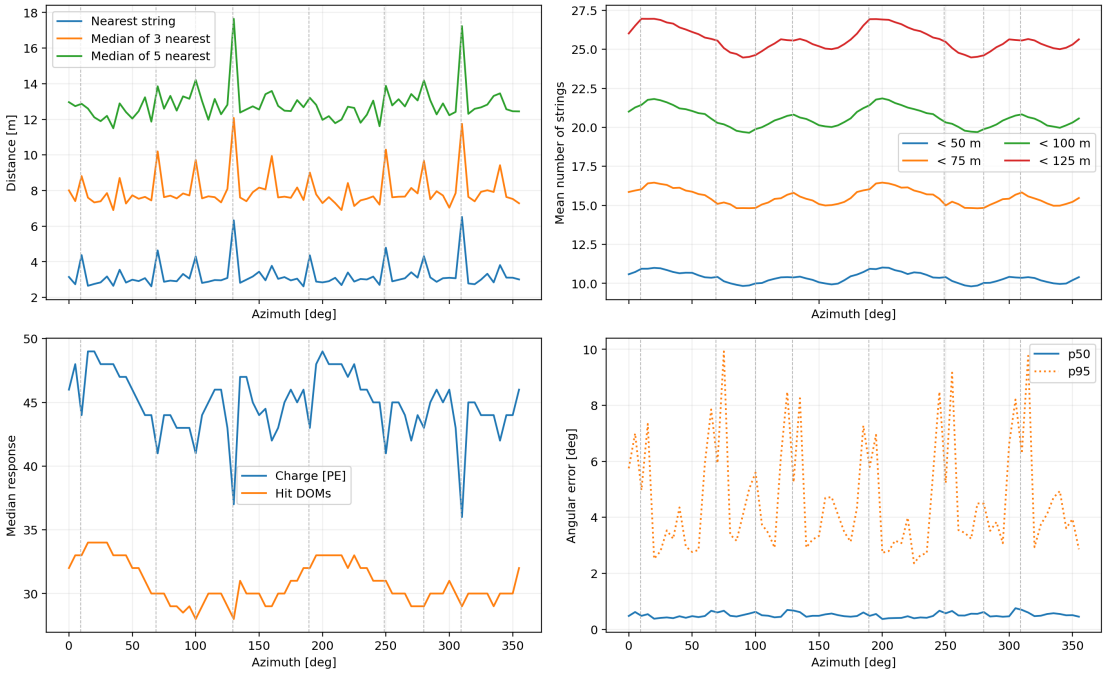


Figure 7.3: Local string-distance metrics, mean numbers of nearby strings, median detected charge and hit-DOM multiplicity, and angular-error percentiles as functions of true azimuth for the pooled final azimuth production. The vertical guides mark the geometrical directions examined in the discussion.

The corridor comparison does not explain every feature of the azimuth scan and is not treated as a complete isolation of all directional effects. It shows that several structures are compatible with the discrete string geometry, whereas the smaller 100°/280° structures are better connected to the event-level distance from nearby strings. The three depth-region productions then test how strongly the surrounding optical region changes the scale of this response.

In clear ice, p_{50} remains between 0.23° and 0.45°. In the dust layer, it ranges from 0.92° to 3.00°, while p_{95} reaches 26.77° and the reconstruction fraction falls to 75.5% in one bin. The same horizontal direction can thus provide a very different constraint depending on the depth region sampled by the event.

The aligned directional bias is ordered in the same way. Its maximum remains 0.165° above the dust and 0.101° in clear ice, but reaches 2.417° in the dust-layer production. In the event-level pooled scan, the maximum is instead 0.662° at 310°, and is dominated by the local zenith component. This aligned result replaces a direct

comparison of the separate mean zenith and azimuth residuals: those coordinate extrema need not occur in the same bin and do not by themselves define one physical angular displacement. The separate depth-region results remain necessary alongside the pooled azimuth curve, but its scale is now expressed by one coherent three-dimensional metric.

7.5 Zenith asymmetry

The nominal zenith production uses 37 bins of 5°, 400 events per bin, uniform azimuths and a perpendicular disk of radius 225 m centred at $(0, 0, -200)$ m (Section 5.4.3). Its directional and spatial results are reported in Figure 6.4 and Table 6.4. This scan introduces a different geometrical difficulty. Azimuth is sampled over the full circle in every bin, so the production already averages over horizontal orientation. The support point, however, is drawn from a disk perpendicular to the track direction. As this disk rotates with zenith, the physical depth range covered by the sampled positions changes as well. In addition, the final bins at θ and $180° - \theta$ were generated independently. The difference between the two sides of the scan could not be assigned directly to the track orientation without a dedicated control.

The starting observation is that the contrast appears in the reconstruction and in the available light at the same time. At 20°, the median angular error is 0.51°, with a median charge of 74 PE and 51 hit DOMs. At 160°, these quantities become 1.00°, 31 PE and 20 DOMs. The poles follow the same ordering: the median errors are 0.88° at 0° and 1.24° at 180°, while the median charges are 53 and 24 PE. The poorer side of the reconstruction is thus also the side that provides less detected information. Figure 7.4 follows this relation over the full zenith range.

The first ambiguity to remove is the independent sampling of the opposite zenith bins. In the paired control, a support point $\mathbf{r}_0$ and a direction $\hat{\mathbf{u}}$ define one infinite line, which is then evaluated again with the opposite orientation $-\hat{\mathbf{u}}$. The support point, closest-approach distances and complete line geometry are unchanged; only the physical direction assigned to the line is reversed. For every direction pair, the control uses 36 azimuths and 100 paired Poisson realisations per line. The ratios in Table 7.2 show that the asymmetry remains under this stricter comparison.

Paired directions	**Expected** Q	**Expected** N_{hit}	**Poisson** Q	**Poisson** N_{hit}
0°/180°	2.13	2.20	2.13	2.20
20°/160°	1.50	1.88	1.49	1.88
45°/135°	1.50	1.75	1.50	1.75
65°/115°	1.34	1.41	1.34	1.41
90°/90°, opposite ϕ	1.00	1.00	1.00	1.00

Table 7.2: Paired-direction light-response control on identical geometrical lines. Ratios use the smaller-zenith orientation divided by its opposite; at 90°, the comparison uses $\phi \to \phi + 180°$.

The contrast is strongest at the poles. Before Poisson sampling, the expected charge ratio is 2.13 for 0°/180°, while the expected ratio of non-empty DOMs is 2.20. The corresponding mean charge difference is 66.86 PE, with a 95% bootstrap interval of $[42.55, 108.77]$ PE. For 20°/160°, the two ratios are 1.50 and 1.88, and the

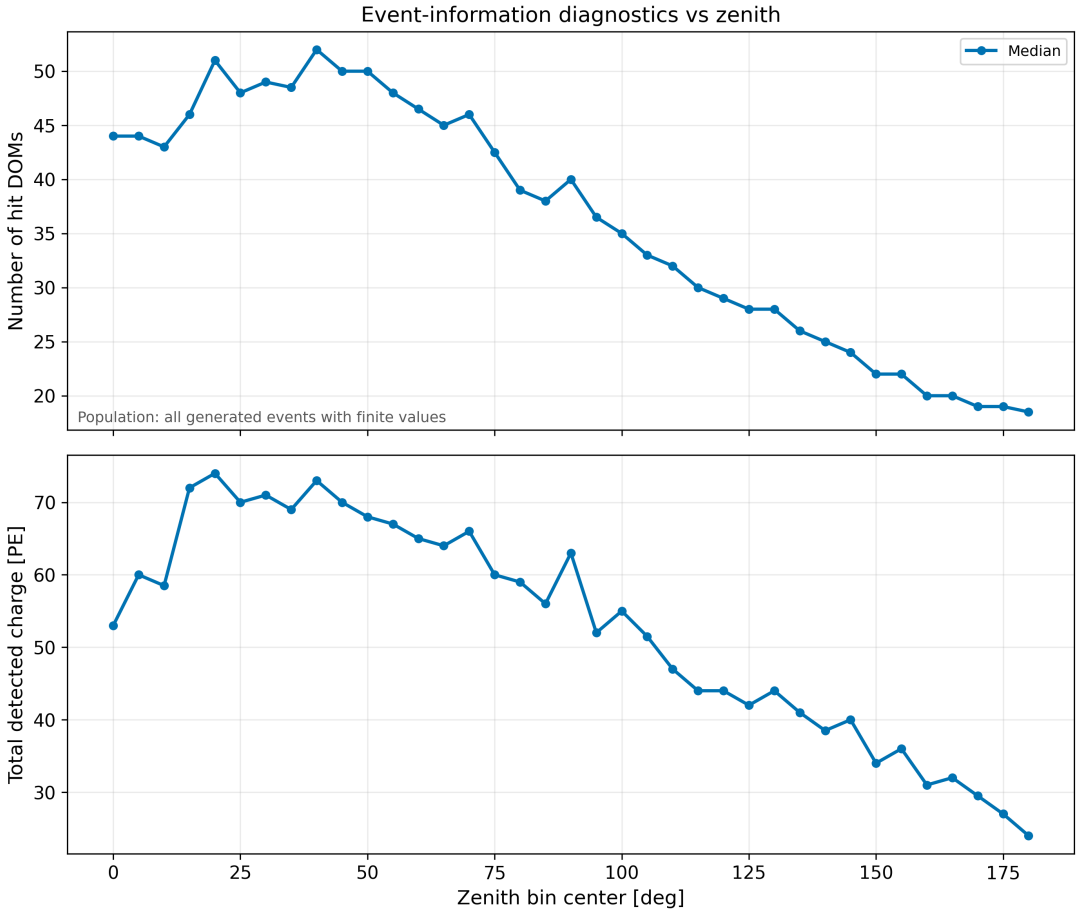


Figure 7.4: Median detected charge and hit-DOM multiplicity as a function of zenith. The figure is used here as a diagnostic of the directional asymmetry shown in Chapter 6.

mean charge difference is 27.54 PE with an interval of $[18.51, 39.46]$ PE. The same ordering remains for $45°/135°$ and $65°/115°$. It disappears for a horizontal $90°$ track compared with the opposite azimuth, where both ratios are close to unity. Reversing the orientation of a line is not enough by itself: no equivalent difference appears for the $90°$ pair. The asymmetry remains in the tested up–down pairs, where the vertical component is reversed. This control isolates the directional component of the effect. The detector design provides the expected physical interpretation below.

The next step is to check whether counting fluctuations create or amplify this ordering. Paired Poisson sampling leaves the ratios essentially unchanged: they remain 2.13 and 2.20 for charge and hit multiplicity at $0°/180°$, and 1.49 and 1.88 at $20°/160°$. The asymmetry is already present in the expected charge response, before the Poisson draw, and cannot be explained solely by the independently sampled track positions of the final production. The small reconstruction check with six events per orientation is not used to infer p_{95} or a reconstructed bias, since its statistics are insufficient for either quantity.

This ordering is expected from the downward-facing PMTs used in the IceCube DOMs [5, 23]. Their angular acceptance favours light arriving towards the downward-oriented photocathode and gives a larger expected response for the tested upgoing tracks. The paired construction should consequently be read as a measurement of how the encoded PMT orientation affects the expected charge on otherwise identical infinite lines, rather than as an unexplained asymmetry. Reversing the muon changes the oriented network inputs even when the geometrical line remains identical, and the trained response carries this angular acceptance. The depth-dependent optical environment and redistribution of informative DOMs can still modulate the magnitude. Because the same learned response is used for generation and reconstruction, this control does not independently validate the size of the PMT angular response in real data.

The remaining sampling question can be examined with the three completed zenith productions. The nominal disk is centred at $z = -200$ m, the second near the geometrical DOM centroid at $z = -23.91$ m, and the third at $z = +200$ m. Figure 7.5 places their main observables on the same axes. The negative and central samples remain relatively close over much of the central zenith range. The upper-centred production instead develops a broader degradation between approximately $105°$ and $155°$, where its charge and hit-DOM curves are also lower. Changing the vertical sampling region changes the reconstruction and the amount and spatial multiplicity of the detected information together. The aligned bias panel gives the same comparison without using separate angular coordinates. Excluding the pole-centred bins, the largest $b_{3\mathrm{D}}$ values are $0.242°$ at $160°$ for the detector-centred sample, $0.329°$ at $170°$ for the nominal -200 m sample and $0.371°$ at $140°$ for the $+200$ m sample. The last value lies inside the intermediate range where the upper-centred production also has poorer information and resolution. This correspondence is descriptive; the different sampled depth distributions still prevent it from identifying one unique cause.

The median describes only the typical part of this change. As shown in Figure 7.6, the $+200$ m sample also develops a distinctly larger p_{95} over the same intermediate angular range, while all three productions become poor near the upper

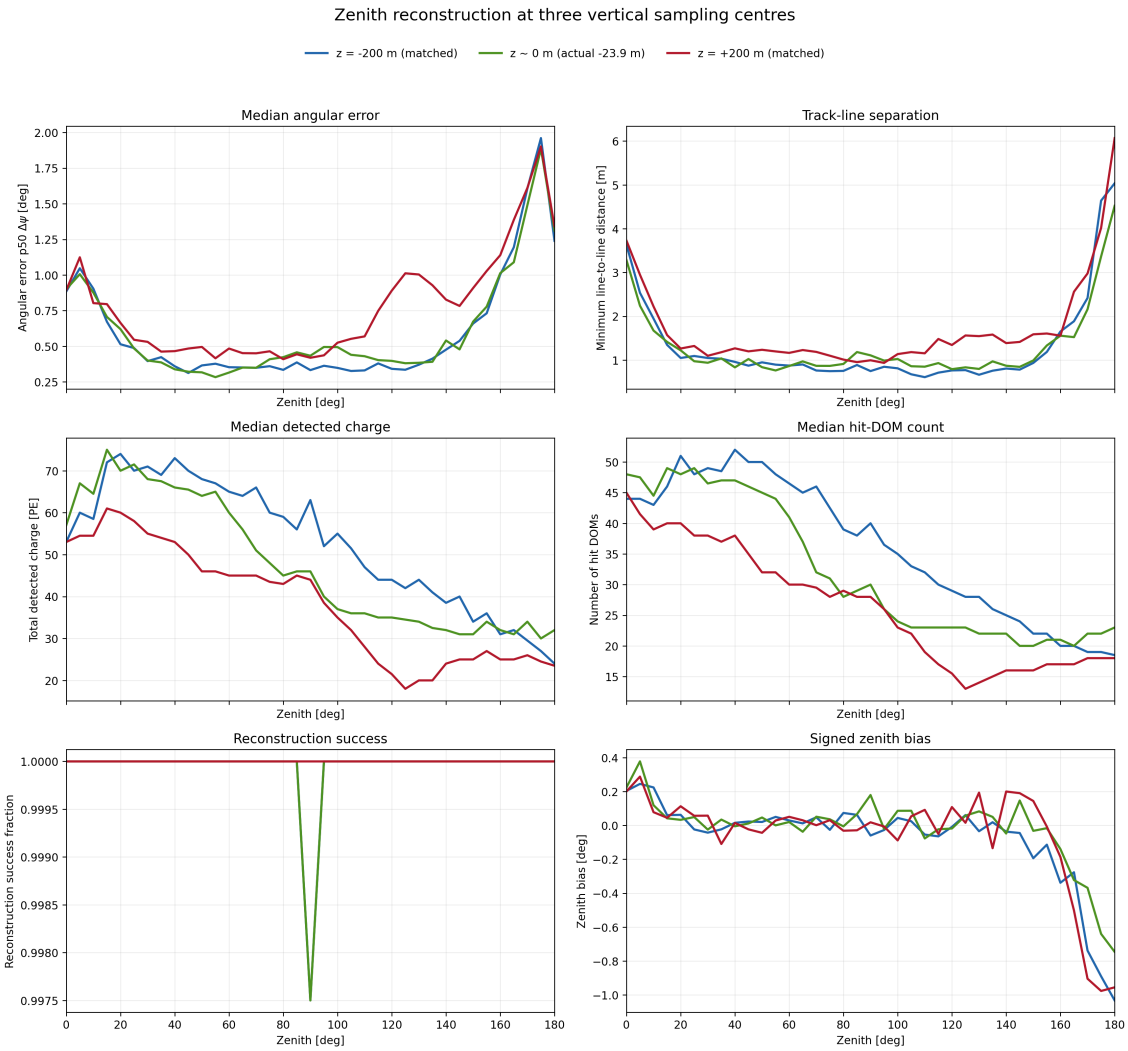


Figure 7.5: Median angular error, line-to-line separation, detected charge, hit-DOM multiplicity, reconstruction success and aligned three-dimensional directional-bias magnitude for the three vertical sampling centres. The $z = -200$ m and $z = +200$ m samples use matched relative track geometries; the detector-centred sample uses a smaller disk and an independent seed. Open markers identify pole-centred bins, whose local component orientation is convention-dependent.

end of the zenith interval. Because rotating the disk changes the physical depth distribution of its tracks, this comparison cannot isolate the dust layer as one unique cause. It establishes instead that the zenith limit depends on the vertical sampling region. It also justifies keeping the −200 m production as the nominal scan, since that reference was deliberately placed in comparatively clear ice.

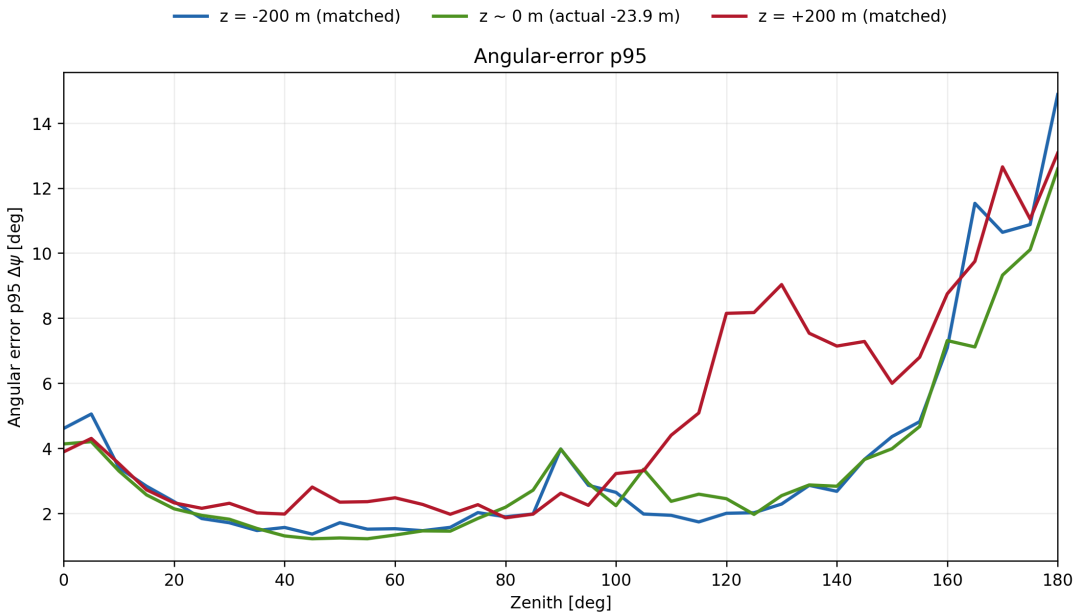


Figure 7.6: Comparison of the zenith angular-error p_{95} for the three vertical sampling centres.

The paired expected-charge sample helps explain why the sampling centre can matter: zenith changes which depths carry most of the optical information. At 20°, the charge-weighted mean DOM coordinate is $z = -125.60$ m and the weighted median is −245.85 m. At 160°, they shift to −344.13 m and −420.67 m. The informative DOM population moves substantially in depth as the direction is reversed, as shown in Figure 7.7. This redistribution is compatible with a contribution from the depth-dependent optical environment, but it still does not isolate one dust or clear-ice layer as the origin of the up/down asymmetry. For this reason, the vertical guides in the figure are used only as reference positions for the representative depth-scan regions.

7.5.1 Coordinate singularity close to the poles

For the usual directional parameterisation

$$\hat{\mathbf{u}}(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\,,$$

the sensitivity of the physical direction to azimuth is

$$\left\| \frac{\partial \hat{\mathbf{u}}}{\partial \phi} \right\| = |\sin\theta|.$$

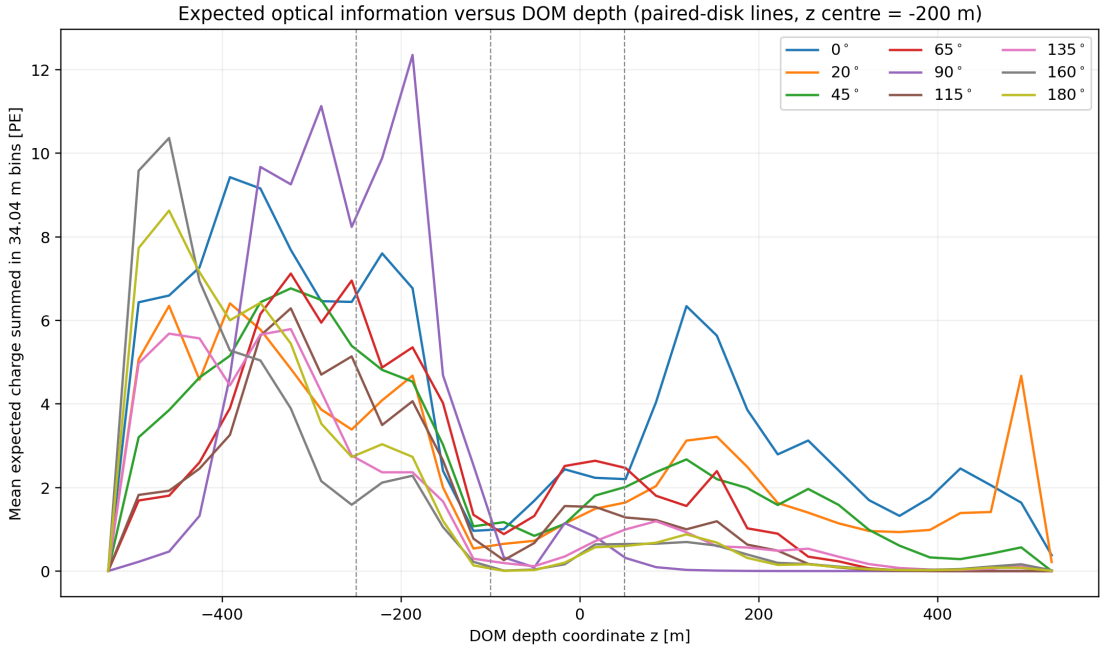


Figure 7.7: Expected charge summed in DOM-depth bins for representative zenith directions in the paired-disk control centred at $z = -200$ m. The dashed vertical guides indicate the representative depth-scan positions only; the figure does not constitute an isolation test of a particular ice layer.

The factor $|\sin\theta|$ tends to zero when $\theta \to 0°$ or $\theta \to 180°$. At the exact poles, changing ϕ leaves the physical direction unchanged, and close to them a broad range of azimuth values corresponds to almost the same track direction.

This coordinate singularity matters when separate zenith and azimuth residuals are averaged near the ends of the scan. A physical azimuthal displacement is proportional to $\sin\theta$, so a raw $\Delta\phi$ can appear large while representing a small change of direction. The boundary of the zenith coordinate also distorts a coordinate mean: close to the north pole, for example, reconstructed zenith values cannot extend below zero. A symmetric directional spread can produce a non-zero mean coordinate residual without a preferred direction in the local tangent plane.

The aligned construction in Section A.1 removes this ambiguity by transporting each true–reconstructed pair to a common local frame before averaging. In the nominal $z = -200$ m production, the largest non-pole value is $b_{3D} = 0.329°$ at 170°. The pole-centred 180° bin reaches 0.673°, but the signed orientation of its local components is fixed by the stated roll convention rather than by an intrinsic azimuth at the pole. Figure B.8 marks these endpoint bins explicitly. The magnitude can be compared with the event-wise angular-error distribution, whereas the pole-component signs are not assigned a separate physical interpretation.

The situation is different at 90°, where $|\sin\theta| = 1$ and a change in azimuth produces the largest directional change. The median angular error is 0.33° at this angle, placing the horizontal configuration among the well-reconstructed central bins. The overall minimum is nevertheless 0.31° at 45°. The horizontal bin is well reconstructed, but it is not the best-performing angle in this production.

7.6 Detector-centred radial dependence and event selection

The detector-centred radial production contains 15 closest-approach radii from 0 to 700 m, 1,000 events at each radius, isotropic directions and random tangent positions around the geometrical DOM centroid (Section 5.4.4). The angular ranges, reference values and event fractions are reported in Figures 6.5 and 6.6 and Table 6.5. After the directional scans, this production asks how far the track can be moved away from the central region before the available constraint changes qualitatively. This distance should not be interpreted as a sharp spherical boundary of IceCube. The detector footprint is irregular and its instrumentation is not spatially uniform. Each radius instead averages a set of isotropic tangent directions with the same closest-approach distance to the detector centre. The resulting curve follows the transition from well-contained tracks to configurations near and beyond the instrumented region.

Two changes have to be followed through this transition. The angular-error percentiles describe only the successfully reconstructed events, whereas the reconstruction fraction keeps all requested events in its denominator. If the percentiles deteriorate while this fraction falls, the surviving events are themselves becoming less precise at the same time as more events are being lost.

In Figure 6.6, the detected and reconstructed curves overlap exactly because every detected event returns a storable fit. The two visible traces represent two coincident pairs, not omitted data.

The scan does not deteriorate immediately. Its smallest median occurs at 150 m, where $p_{50} = 0.41°$, and the median remains below 1° through 400 m. It first exceeds this value at 450 m and reaches 1.61° at 500 m. The stronger low-information regime begins farther out. At 600 m, $p_{50} = 5.40°$, $p_{95} = 37.65°$ and $f_{\mathrm{reco}} = 93.0\%$; by 700 m, they have reached 12.74°, 59.30° and 58.2%. Over the same range, the median hit multiplicity and charge fall from 33 DOMs and 50 PE at 150 m to one DOM and 1 PE at 700 m. A coherent directional effect also appears in this outer regime: $b_{3\mathrm{D}}$ reaches 3.13° at 700 m, with an interval of $[1.56, 4.86]°$. Its local zenith component is +3.13°, while the local azimuthal-arc component remains compatible with zero. The former raw mean zenith residual of +3.80° overstated the coordinate-independent magnitude, but the aligned calculation confirms that the outermost reconstructed population retains a coherent shift dominated by the local zenith direction.

The expected-charge convention also has to be kept in this interpretation. Responses are set to zero for a DOM more than 400 m from the true track, whereas the radial coordinate is the distance from the track to the detector centroid. The two quantities are not the same, so the scan does not contain an imposed discontinuity at $R = 400$ m. The per-DOM cutoff can nevertheless reduce the available light for distant tracks and is one of the conditions under which the outer degradation is measured.

The shallow minimum at 150 m does not mean that moving away from the detector centre intrinsically improves the reconstruction. Relative to the central bin, it coincides with an increase from 26 to 33 median hit DOMs and from 40 to 50 PE. Each radius samples a different family of tangent lines through the discrete string

geometry, so the 150 m population can be better illuminated than the nominally central one. This is consistent with the placement dependence found in the exploratory scans. Only at larger radii does the loss of information become the dominant and nearly monotonic effect.

The area-weighted integration gives a compact reference for this dependence without discarding the shape of the scan. For isotropic directions and closest-approach points uniform within $R_{\mathrm{ref}} = 500$ m, the mean angular error is 1.72°, while $p_{50} = 0.74°$, $p_{68} = 1.32°$ and $p_{95} = 6.32°$. The reconstruction fraction is 99.97% and the area-weighted 3D bias of this reference population is 0.13°, with a bootstrap standard deviation of 0.03°. The mean is larger than the median because a minority of less favourable events already contributes to the upper tail. This is the main reference value of the study, but it is tied to the declared 500 m acceptance and to the common 100 TeV-motivated light-yield setting.

Including the complete range to 700 m changes the reference to a more severe population: the mean becomes 5.76°, the median 1.47°, $p_{95} = 27.46°$ and the reconstruction fraction 93.03%. These values do not contradict the 500 m result. They demonstrate directly that a quoted angular value depends on which outer tracks are admitted and that the associated event-survival fraction must be stated with it.

Extending the closest approach beyond the instrumented volume also turns this production into a stress test of the model. An infinite track may still illuminate part of the array there, but the remaining geometrical constraint becomes weak. Repeating the scan at several light-yield scales could later determine how far outside the instrumented volume a track remains reconstructable for a chosen angular-precision requirement. This last point is an outlook rather than a result of the present production.

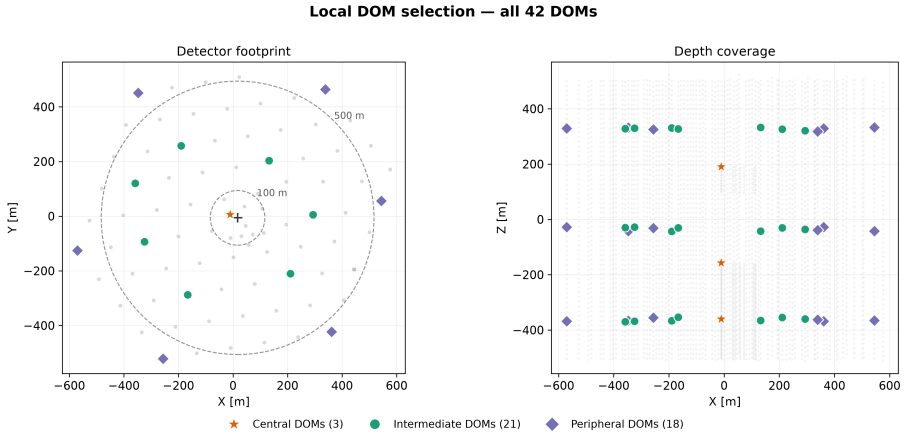
7.7 Local DOM populations and the origin of the large tails

The primary local production uses 24 selected DOMs, of which 18 are outer, and six fractional neighbour distances. Tracks are nearly horizontal, $\theta \in [89°, 91°]$, and each direction is repeated through three independent light realisations (Section 5.4.5). The radial and above–below results are shown in Figures 6.7 and 6.8, with their upper-tail and bias summary in Table 6.7. This production brings the discussion back to the local scale. It was designed around two questions: how the reconstruction changes with transverse distance from a selected DOM, and whether the two sides of that DOM behave differently. The curves answer these questions, but their large p_{95} values cannot be understood from the local coordinates alone. Increasing the transverse distance changes the typical reconstruction only modestly and never removes the broad upper tail. In the vertical scan, the ordering of the above- and below-DOM placements reverses after the closest offsets instead of revealing one side that is always favourable.

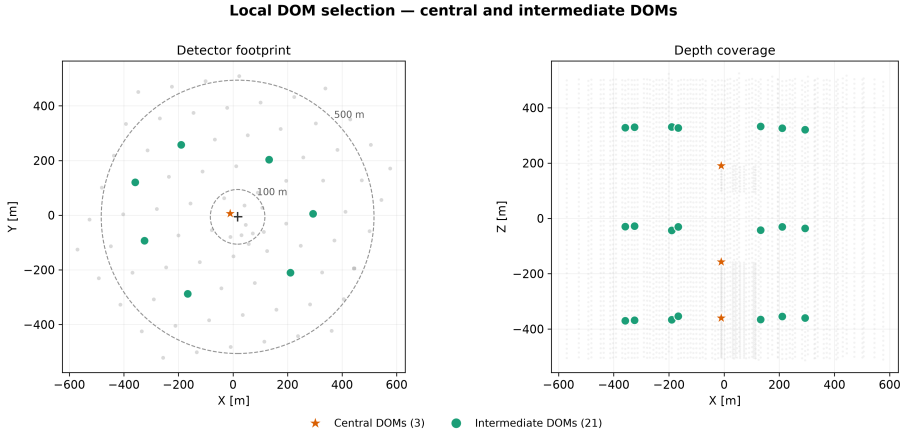
The composition of this first sample gives a possible explanation. Although it averages over 24 locations, 18 of them are outer DOMs, while only 3 are intermediate

and 3 central. The production is strongly weighted towards the detector periphery. An extended 42-DOM sample was added to broaden the spatial coverage and, more importantly, to test whether this detector position is responsible for the large tails.

The extended sample contains 3 central, 21 intermediate and 18 outer DOMs. By retaining only its central and intermediate locations, the same reconstructed events can also be summarised as a 24-DOM subset. This is not the primary 24-DOM production of Chapter 6: it contains different modules and serves only as a regional sensitivity check. Figure 7.8 shows the extended 42-DOM population (Figure 7.8a) and the corresponding 24-DOM subset (Figure 7.8b) used in this diagnostic.



(a) Extended 42-DOM selection.



(b) Centre + intermediate 24-DOM subset.

Figure 7.8: DOM populations used to test the spatial origin of the upper angular-error tails.

The first comparison is made directly on the angular-error curves in Figures 7.9 and 7.10. Removing the outer locations has only a limited effect on the median, but a much larger one on p_{95}. The full 42-DOM average still contains a broad tail, whereas its centre + intermediate subset is consistently narrower in both local scans.

Event survival remains almost complete in these comparisons. In the radial scan, the extended 42-DOM sample reaches a minimum reconstruction fraction of 98.76% in the outermost bin, whereas the centre + intermediate subset remains at 100% throughout. Both vertical populations are fully reconstructed at every tested offset.

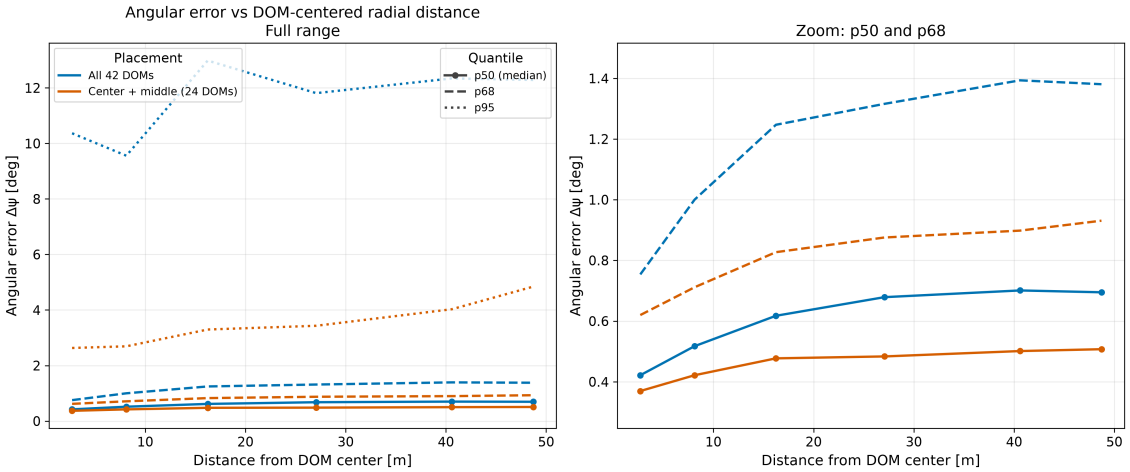


Figure 7.9: Radial angular-error comparison between the extended 42-DOM selection and its centre + intermediate subset.

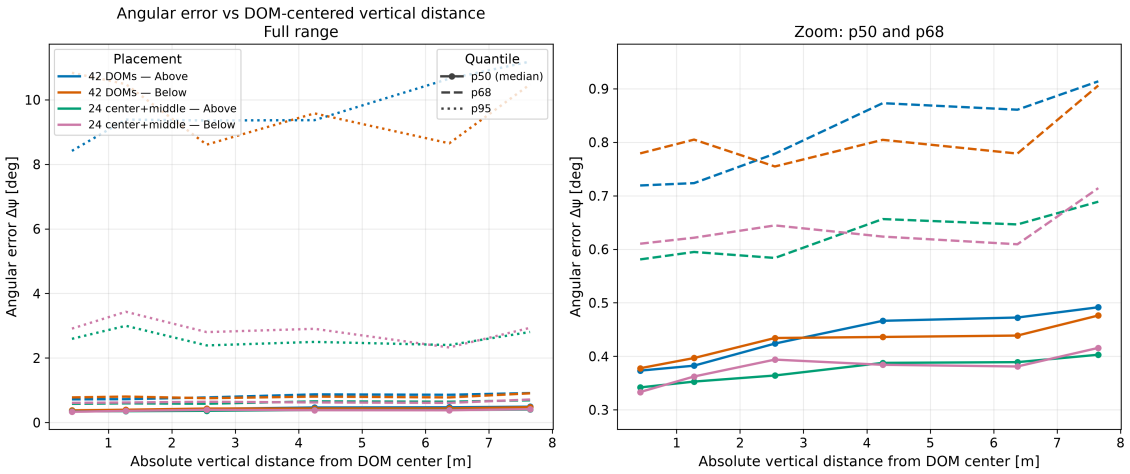


Figure 7.10: Vertical angular-error comparison between the extended 42-DOM selection and its centre + intermediate subset.

The population dependence is mainly visible in the quality and upper tail of the reconstructed events rather than through a loss of events.

To avoid mixing the two different 24-DOM selections, their numerical summaries are kept separate in Table 7.3. Broadening the initial sample to 42 DOMs already reduces the mean tail relative to the edge-weighted primary production. The larger reduction appears only after the outer group is removed from the extended sample, and the maximum absolute mean zenith residual follows the same ordering.

Scan	**Population**	**Mean** p_{95}	**Max** p_{95}	**Max** $\lvert\langle\Delta\theta\rangle\rvert$	**Max** $\lvert\langle\Delta\phi\rangle\rvert$
Radial	Primary 24	19.54	20.49	1.42	0.20
Radial	Extended 42	11.56	12.98	0.82	0.12
Radial	Centre + intermediate 24	3.49	4.84	0.07	0.08
Vertical	Primary 24	18.73	23.15	1.71	0.38
Vertical	Extended 42	9.75	11.18	1.00	0.22
Vertical	Centre + intermediate 24	2.75	3.43	0.05	0.05

Table 7.3: Comparison of the primary local production and the two populations derived from the extended diagnostic. All angular quantities are in degrees. The last two columns contain bin-level mean coordinate residuals, not components of the aligned directional bias.

The regional split of the 42-DOM sample tests this explanation more directly. Table 7.4 shows a continuous progression from central to intermediate and then outer locations. In the central group, p_{95} remains close to one degree and no event exceeds 10°. In the outer group, it reaches about 24.4° in both local scans, with more than 10% of events above 10°. Charge and hit multiplicity decrease along the same progression, although neither quantity alone determines the error of an individual event.

Scan	**Region**	p_{50}	p_{68}	p_{95}	$> 10°$	$\langle\Delta\theta\rangle$	**Q** / N_{hit}
Radial	Central	0.26	0.40	1.24	0.00%	+0.00	68 / 41
Radial	Intermediate	0.50	0.88	3.75	0.68%	+0.03	34 / 19
Radial	Outer	0.96	2.32	24.40	12.70%	+1.24	26 / 15
Vertical	Central	0.24	0.37	1.20	0.00%	-0.01	91 / 42
Vertical	Intermediate	0.40	0.69	2.98	0.47%	+0.01	50 / 20
Vertical	Outer	0.55	1.41	24.72	10.89%	+1.70	39 / 16

Table 7.4: Local reconstruction metrics after separating the extended 42-DOM sample into central, intermediate and outer regions. Angular quantities are in degrees; $\langle\Delta\theta\rangle$ is the mean signed zenith coordinate residual.

The DOM-level values in Figure 7.11 preserve the same separation. One central DOM contributes more than the other two to the small tail of that category. However, removing the central DOMs one at a time leaves the pooled central p_{95} between 0.89° and 1.50°. Occasional unfavourable events can still occur near the detector centre, but none of the central DOMs approaches the outer population. Position need not be the only cause at event level; it is the dominant population-level explanation isolated by this comparison.

The signed coordinate residuals are also compared at DOM level. As in the primary local curves, the estimator is the mean of the per-DOM means, with uncertainty bands obtained from 2,000 DOM-level bootstrap resamples. The comparison changes the selected DOM population, not the resampling unit. When the outer

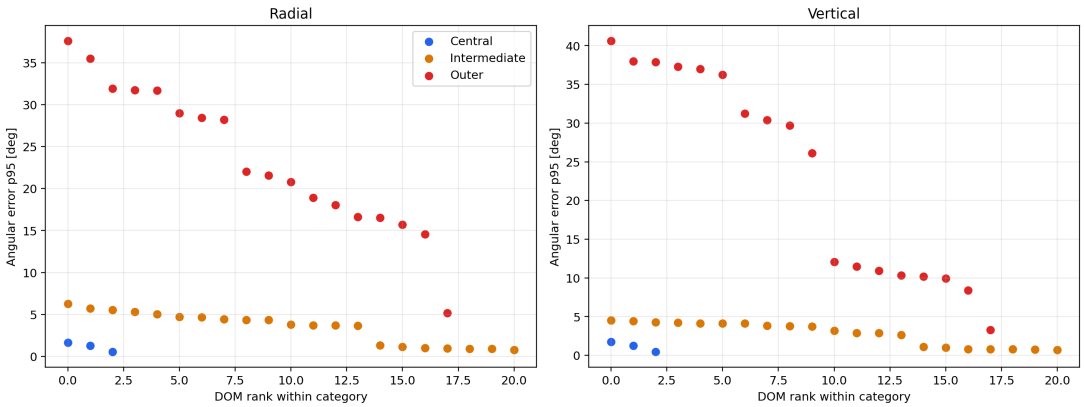


Figure 7.11: Per-DOM p_{95} values for the radial and vertical local-placement diagnostics, separated into central, intermediate and outer detector regions. DOMs are ranked within each category for readability.

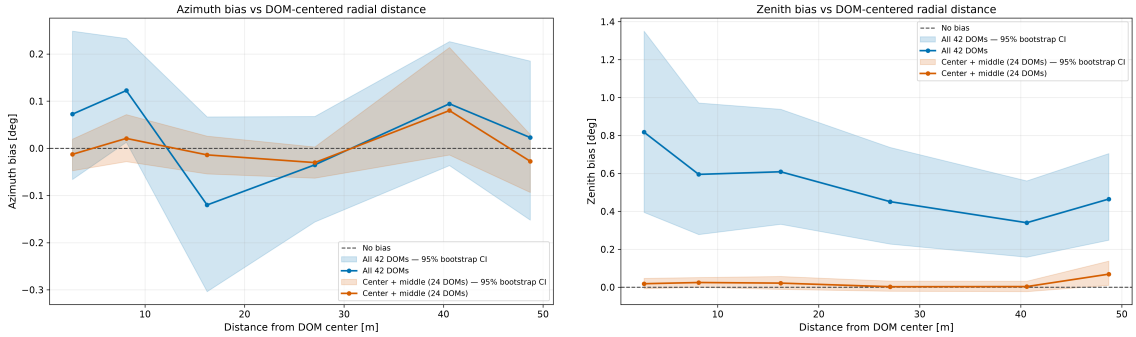
group is removed, the maximum absolute mean zenith residual falls from 0.82° to 0.07° in the radial scan and from 1.00° to 0.05° in the vertical scan. The separate zenith-residual intervals do not overlap in any comparison bin. The 24-DOM population is a subset of the 42-DOM sample, however, so these intervals do not describe a paired difference and cannot be used as a formal significance test. The azimuth-residual intervals overlap in every bin, and no equivalent statement is supported for that component.

The corresponding radial and vertical signed-coordinate-residual comparisons are shown in Figures 7.12 and 7.13, respectively.

The common positive mean zenith residual in the primary above–below plot is consistent with this population dependence. Moving a support point above or below a DOM does not reverse the track, and both sides use the same near-horizontal zenith distribution. Their residuals are not required to have opposite signs. In the regional split of Table 7.4, the vertical zenith mean is −0.01° for central DOMs, +0.01° for intermediate DOMs and +1.70° for outer DOMs. The positive component in the edge-weighted primary average is consequently driven mainly by its outer-DOM population. It should not be interpreted as the above–below effect itself; that effect is the difference between the two curves at the same offset. The azimuth point estimates have a mirrored appearance, but their bootstrap intervals include zero and do not resolve a non-zero mean azimuth residual.

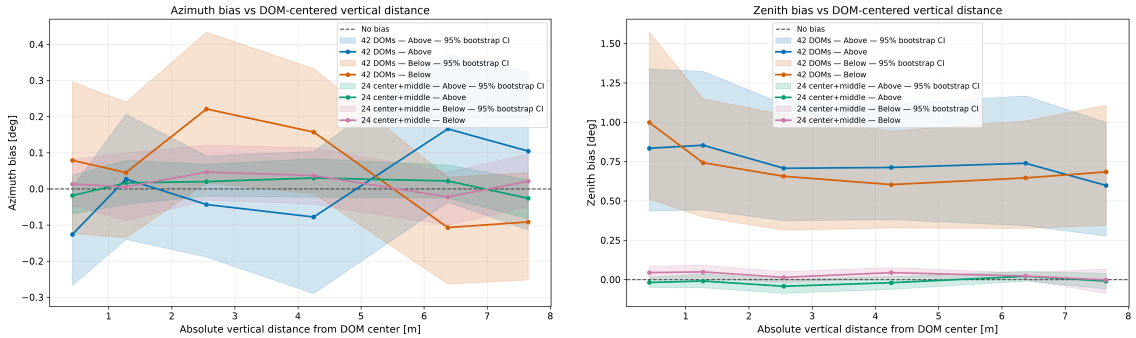
The expected downward-facing PMT acceptance motivates the original above–below question and makes the scan a measurement of how that orientation is expressed in the matched detector response. The measured reconstruction curves still combine this acceptance with the surrounding string geometry, depth-dependent optical conditions and detector containment, so they do not isolate the orientation contribution quantitatively. In particular, the reversal of the above–below ordering with offset shows that one side is not systematically better in this edge-weighted reconstruction sample. The two closest offsets also show that total charge alone does not determine the local reconstruction quality. Below the DOM, the median charge reaches 99 PE instead of 46 PE at 0.425 m, and 65 PE instead of 39 PE at 1.275 m. The angular-error percentiles are nevertheless slightly worse on that side, while the

median number of hit DOMs decreases from 19 to 18. In these two bins, the side receiving more total charge is not the side providing the better reconstruction. The additional charge may be concentrated among a less favourable set of DOMs and may carry less useful spatial or timing information. The scan does not distinguish between these possibilities. A quantitative isolation of the PMT-orientation contribution would require another control in which the angular acceptance is changed while the surrounding geometry and optical conditions are held fixed.



(a) Mean azimuth-residual comparison. **(b)** Mean zenith-residual comparison.

Figure 7.12: Radial signed-coordinate-residual comparison between the extended 42-DOM selection and its centre + intermediate subset.



(a) Mean azimuth-residual comparison. **(b)** Mean zenith-residual comparison.

Figure 7.13: Vertical signed-coordinate-residual comparison between the extended 42-DOM selection and its centre + intermediate subset.

7.8 Directional bias, tails and statistical interpretation

Bringing the scans together shows why no single metric is sufficient to describe the reconstruction. The comparisons in this section retain the scan populations defined in Table 5.1 and the local populations defined in Section 5.4.5; no new simulation is introduced. The depth production already gives one possible combination: p_{95} exceeds 10°, yet the bootstrap intervals of the signed-coordinate-residual extrema include zero (Figure 6.1 and Table 6.1). The reconstruction becomes less precise

because the distribution broadens, without a resolved displacement towards one direction.

The local scans separate the typical event from the tail in another way. In the extended 42-DOM production, both medians remain below one degree while p_{95} exceeds 11°. Removing the outer locations makes the tails much smaller, although the median and upper part of the distribution still do not carry the same information (Table 7.3). The clear-ice azimuth production shows the same distinction on a smaller scale.

When the optical and geometrical constraint becomes still weaker, broad errors can appear together with event loss. This is clearest in the detector-centred radial scan. At 700 m, $p_{50} = 12.74°$ and $p_{95} = 59.30°$, only 58.2% of the requested events are reconstructed, and the aligned bias reaches $b_{3D} = 3.13°$, dominated by a positive local zenith component (Table 6.5 and Figure 6.6). The dust-layer azimuth scan is an intermediate case: most events are retained, but one bin reaches a minimum reconstruction fraction of 75.5% while the scan also contains a heavy tail and a larger coherent directional shift.

Since these behaviours can appear separately or together, p_{50}, p_{68}, p_{95}, b_{3D}, its signed local components and the event fractions must remain separate observables. The area-weighted mean is a useful reference, but using it alone would hide the distinction between a change in the typical reconstruction, a minority tail, a directional shift and a loss of events.

The bootstrap intervals have a narrower role within this set of metrics. They quantify the sampling uncertainty of the chosen bin-level directional estimator. They are neither the p_{68} or p_{95} spread of individual reconstruction errors nor a per-event angular-uncertainty estimate. Non-parametric bootstrap resampling is a standard way of estimating the sampling distribution of a statistic [41]. More generally, confidence-region statements depend on the statistical construction and the associated test statistic; this distinction is important in neutrino analyses and is reviewed, for example, by Agostini and Neumair [42]. The present bootstrap is used only to determine how precisely the bin-level bias has been estimated. In particular, the percentile interval of the non-negative b_{3D} magnitude is not treated as a formal test of zero bias; the signed local components are retained to describe the direction of the shift.

For the local study, the population comparison in Section 7.7 supplies the interpretation of the tail: peripheral locations dominate the large values, while the central and intermediate populations remain narrower. This conclusion applies at population level and does not exclude rare unfavourable events in the detector centre.

7.9 Charge, hit multiplicity and reconstruction precision

Across the cases discussed above, total charge indicates how much optical information is available, but it does not predict the angular error on its own. This comparison uses the detected-information panels already introduced for depth (Figure 7.1),

zenith (Figure 7.4) and radius (Table 6.5); it does not define a new event sample. The number of hit DOMs, their spatial distribution and the photon timing information determine how useful that charge is for the fit.

Separating the azimuth scan into the three depth regions provides the clearest scale comparison. Clear ice gives median charges of 67–87 PE and 46.5–59 hit DOMs, whereas the dust layer gives only 4–12 PE and 3–6 hit DOMs. Its reconstruction is accordingly much worse, although the angular error does not change in direct proportion to either quantity. The zenith scan gives the same connection within one production: from 20° to 160°, the median charge falls from 74 to 31 PE while the median angular error rises from 0.51° to 1.00°. The paired-direction control makes this observation more specific by showing that the hierarchy in charge and hit-DOM multiplicity is already present before Poisson sampling on identical geometrical lines. Near the detector boundary, the relation becomes more extreme. At 700 m, the sample reaches one median hit DOM, 1 PE and $f_{\mathrm{reco}} = 58.2\%$, together with poor angular precision.

These comparisons do not define a universal charge threshold. A track must produce enough light in a geometrically informative set of DOMs, and the same amount of charge can carry a different directional constraint depending on where and when it is detected. A dedicated correlation or conditional analysis could quantify this relation further, but it is not required to support the interpretation made here.

7.10 What is established by the present study

The four production scans in Figures 6.1, 6.2, 6.4 and 6.5, together with the local population comparison in Table 7.4, establish that the reconstruction limit remains configuration dependent even in the matched-model setup. Depth, direction, distance to the detector centre and the local DOM population all change the error distributions while generation and reconstruction continue to use the same underlying model. The main variations can be connected to three detector-dependent ingredients already encountered throughout the chapter: the depth-dependent optical profile, the local instrumentation geometry and containment within the instrumented volume.

Sampling the nuisance variables is still necessary because it prevents one accidental local configuration from defining an entire result. It does not, however, remove these physical dependencies. The azimuth results separated by depth region and the comparison of the local DOM populations make this explicit. The dust-rich region is consistently less favourable than the tested clearer regions, with larger angular errors, less light and fewer hit DOMs. Several azimuthal structures are also numerically compatible with the principal string corridors, although this comparison does not explain every feature of the scan.

The zenith and radial controls complete the picture in different ways. The up/down light-response asymmetry remains when opposite orientations are tested on exactly the same infinite geometrical line, and paired Poisson sampling preserves the deterministic ordering. Its sign is consistent with the expected angular acceptance of the downward-facing PMTs. The difference cannot come from independently

sampled positions because the paired directions share the same geometrical line. It is already present in the expected charge response, before the Poisson draw and before reconstruction. The optimiser is not required to create it. The fit can nevertheless change how strongly the light-response difference appears in the final angular errors. The three vertical sampling centres additionally show that the angular range and magnitude of the degradation depend on the region around which the rotating disk is constructed. At local scale, separating the DOM categories places the large tails mainly in the outer population. At detector scale, moving the closest approach outward eventually reaches a low-information regime in which angular precision and event survival deteriorate together. These are the practical matched-model limits established here. The area-weighted radial integration summarises one declared population, while the separate scans show the range of conditions that this reference does not contain by itself.

7.11 Scope and limitations

These conclusions remain conditional on the idealised matched-model setup defined in Sections 5.1 and 5.1.3, with the production populations listed in Table 5.1. The study introduces no ice-model mismatch, calibration error or intentionally incorrect detector response. It also omits a realistic stochastic muon-loss pattern. The resulting population is not an experimental angular-resolution sample. The area-weighted reference remains valid for the defined isotropic population, radial acceptance and light-yield convention, but cannot be transferred to an arbitrary IceCube event selection.

The aggregated and local populations require the same care. The pooled azimuth curve gives equal requested weight to three representative depth layers instead of following the natural detector population. The primary 24-DOM production is weighted towards the detector edge. The extended 42-DOM diagnostic intentionally keeps 18 outer DOMs, and its centre + intermediate subset is only a regional sensitivity check. None of these local selections is an unbiased detector-wide baseline.

The three-centre zenith comparison also requires care. The -200 m and $+200$ m samples use matched relative geometries, but the detector-centred sample has a different disk radius and an independent seed. Together they establish that the vertical sampling region matters; they do not isolate one ice layer as the unique cause of the observed differences.

This study examines favourable closure conditions rather than a complete event. Stochastic losses matter for high-energy IceCube track reconstruction and motivate dedicated hypotheses [6, 19]. Including them is the next step.

Chapter 8

Conclusion

The objective of this thesis was to determine how accurately and robustly the Triple-Pandel infinite-muon model can reconstruct a muon track in IceCube when the continuous light emission is generated and reconstructed under the same controlled assumptions. The work was also intended to identify which optical, geometrical and sampling conditions structure the remaining reconstruction errors. The study now provides both a well-defined angular reference and the range of detector conditions over which the reconstruction departs from that reference.

The main conclusions are the following.

1. For isotropic directions and closest-approach points distributed uniformly in transverse area within $R_{\mathrm{ref}} = 500$ m, the mean angular reconstruction error is 1.72°, with a 95% bootstrap interval of $[1.65, 1.79]^\circ$. The same population has $p_{50} = 0.74^\circ$, $p_{68} = 1.32^\circ$, $p_{95} = 6.32^\circ$, an area-weighted 3D bias of 0.13°, with a bootstrap standard deviation of 0.03°, and a reconstruction fraction of 99.97%. This is the principal reference value of the thesis. It applies to the matched infinite-muon model with the $s_E = 1.425$ light-yield convention motivated at 100 TeV; it is not an experimental IceCube resolution for an explicitly propagated 100 TeV muon sample. Extending the acceptance to the full 700 m stress-test range changes the population and raises the mean to 5.76°, while the reconstruction fraction falls to 93.03%.

2. The depth-dependent optical environment is one of the main limits on the available directional information. The best depth bin at $z = -364$ m reaches $p_{50} = 0.19^\circ$ and $p_{95} = 0.66^\circ$, whereas the dust-rich region reaches $p_{50} = 1.45^\circ$ at $z = -75$ m and $p_{95} = 10.69^\circ$ at $z = -109$ m. The azimuth results show the same ordering when separated by depth region: the clear-ice median remains between 0.23° and 0.45°, while it ranges from 0.92° to 3.00° in the dust layer. The maximum aligned directional bias also rises to 2.417° in the dust-layer production, compared with 0.165° above it and 0.101° in clear ice. These changes occur together with lower detected charge and fewer hit DOMs. They quantify how the optical profile contained in the common model affects the reconstruction; they are not an independent measurement of the ice properties.

3. Direction and detector geometry remain relevant after averaging over local track positions. Across the pooled azimuth scan, p_{50} changes by 0.38°, from

0.37° to 0.75°, and p_{95} changes by 7.57°, from 2.36° to 9.93°. The upper structures are compatible with the principal string corridors, although these spans do not isolate a corridor-only contribution. Smaller features near 100° and 280° coincide with larger event-level distances from nearby strings. In the zenith study, the paired control shows that the up–down ordering of charge and hit-DOM multiplicity is already present in the expected light response on identical infinite geometrical lines. Its sign is consistent with the downward-facing PMT orientation. The comparison of three vertical sampling centres further shows that the magnitude and angular range of the degradation depend on which detector region provides the information. These controls support the geometrical and detector-response interpretations without claiming that every directional feature has been isolated to one cause.

4. Containment determines both the broadest angular tails and the loss of events. In the detector-centred scan, the median remains below 1° through 400 m, but reaches 12.74° at 700 m, where $p_{95} = 59.30°$ and only 58.2% of the requested events are reconstructed. The local DOM diagnostics lead to the same conclusion from another scale. Central locations retain p_{95} values close to 1.2°, with no event above 10°, whereas outer locations reach about 24.4° and place more than 10% of their events above 10°. The outer detector population, rather than the local offset alone, is the main source of the large tails in the initial DOM-centred production.

These results also show why the mean, percentiles, directional bias and event fractions have to be reported together. The mean gives a useful reference for a declared population, but it does not distinguish a broad minority tail from a coherent displacement of the reconstructed directions. A percentile can in turn describe only the events that were successfully reconstructed and can hide a simultaneous loss of events. The aligned three-dimensional bias avoids the coordinate dependence of separate zenith and azimuth means, particularly close to the poles. Total charge is similarly informative without being sufficient on its own: the number, spatial distribution and timing of the hit DOMs determine how strongly the track is constrained. No universal charge threshold separates good and poor reconstructions in the samples studied here.

Within this matched synthetic framework, TriplePandel is effective at recovering the muon direction for well-illuminated and sufficiently contained tracks. Here, *effective* refers to the angular accuracy, controlled bias and high reconstruction fraction observed in those conditions. It does not mean that TriplePandel has been shown to be computationally faster or more accurate than SplineMPE or SegmentedSplineReco. No same-event comparison of accuracy, execution time or memory use was performed in this work, and the results cannot yet be transferred directly to real IceCube data.

The methodological contribution is as important as the numerical values. The first fixed scans showed how easily one local track placement can be mistaken for a detector-wide behaviour. Bin-wise randomisation, the separation of global and local populations, paired controls and the retention of event-level outputs made it possible to test the initial explanations and, when necessary, to revise them without repeating the simulations. The same event-level information also allowed the directional bias

to be reformulated as a coordinate-independent three-dimensional quantity and the radial scan to be integrated into a reproducible isotropic reference.

The interpretation remains limited by the purposefully favourable closure conditions. Generation and reconstruction use the same optical and detector response, the baseline does not contain realistic stochastic muon losses, noise, calibration errors or an intentional model mismatch, and the neutrino–muon production angle is not reconstructed. The common light-yield factor is motivated by the nominal 100 TeV configuration, but it does not replace the propagation of a complete 100 TeV muon with its individual energy losses. The reconstruction also starts from favourable seeds and is not a test of arbitrary global initialisation. These limitations define what the measured closure level means; they do not remove the detector-dependent effects found within that framework. The outer radial behaviour is also conditional on the implemented 400 m per-DOM expected-charge cutoff, which is not a physical detector boundary.

The next step should be to introduce realistic stochastic-loss tracks and then controlled differences between the generation and reconstruction models while preserving the same event geometries. A same-event comparison with SplineMPE and SegmentedSplineReco should follow, using angular percentiles, aligned bias, reconstruction success, execution time and memory consumption together. Only after these tests would a comparison with realistic Monte Carlo or data provide a meaningful measure of experimental performance. The present work supplies the controlled benchmark required for that progression: it identifies where the continuous-emission, infinite-muon reconstruction already performs well, where detector information limits it, and which improvements must be tested before gains in source-direction reconstruction can be claimed.

Appendix A

Reconstruction Metrics and Statistical Summaries

This appendix clarifies the quantities that are needed to read the results. It does not repeat the event-generation or fitting sequence, which is described in Appendix C. Angles are calculated internally in radians and reported in degrees; positions and distances are in metres, times in nanoseconds, and charge is a simulated photoelectron-count proxy.

A.1 Directional performance

The principal performance quantity is the three-dimensional opening angle between the true and reconstructed directions,

$$\Delta\psi = \arccos(\mathrm{clip}\left[\hat{\mathbf{u}}_{\mathrm{true}}\cdot\hat{\mathbf{u}}_{\mathrm{reco}}, -1, 1\right]). \tag{A.1}$$

It is non-negative and measures the magnitude of the directional discrepancy. The clipping only protects the calculation against numerical round-off outside the domain of the inverse cosine.

Separate coordinate residuals retain information that the opening angle cannot show. Their convention is reconstructed minus true,

$$\Delta\theta = \theta_{\mathrm{reco}} - \theta_{\mathrm{true}}, \qquad \Delta\phi = \mathrm{atan2}[\sin(\phi_{\mathrm{reco}} - \phi_{\mathrm{true}}), \cos(\phi_{\mathrm{reco}} - \phi_{\mathrm{true}})]\,. \tag{A.2}$$

The wrapped azimuth residual avoids an artificial discontinuity at $0°/360°$. These quantities remain useful diagnostics, but their means depend on the angular coordinates. In particular, a change in azimuth represents a physical arc length proportional to $\sin\theta$, and azimuth is not a stable coordinate close to the zenith poles.

The final azimuth, zenith and detector-centred radial scans use an aligned three-dimensional directional bias. For event i, let θ_i and ϕ_i be the true zenith and azimuth, and define the local tangent vectors

$$\mathbf{e}_{\theta,i} = \begin{pmatrix} \cos\theta_i\cos\phi_i \\ \cos\theta_i\sin\phi_i \\ -\sin\theta_i \end{pmatrix}, \qquad \mathbf{e}_{\phi,i} = \begin{pmatrix} -\sin\phi_i \\ \cos\phi_i \\ 0 \end{pmatrix}. \tag{A.3}$$

The reconstructed direction expressed in this roll-fixed local frame is

$$\mathbf{v}_i = \begin{pmatrix} \mathbf{e}_{\theta,i}\cdot\hat{\mathbf{u}}_{\mathrm{reco},i} \\ \mathbf{e}_{\phi,i}\cdot\hat{\mathbf{u}}_{\mathrm{reco},i} \\ \hat{\mathbf{u}}_{\mathrm{true},i}\cdot\hat{\mathbf{u}}_{\mathrm{reco},i} \end{pmatrix} = R_y(-\theta_i)R_z(-\phi_i)\hat{\mathbf{u}}_{\mathrm{reco},i}. \tag{A.4}$$

The same rotation maps every true direction onto $(0,0,1)^{\mathsf{T}}$. The vectors $\mathbf{v}_i$ can then be averaged inside a bin without mixing different true directions. For N successful reconstructions, define

$$\mathbf{m} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{v}_i, \qquad \hat{\mathbf{m}} = \frac{\mathbf{m}}{\|\mathbf{m}\|}, \qquad t = \sqrt{\hat{m}_x^2 + \hat{m}_y^2}. \tag{A.5}$$

The aligned three-dimensional bias magnitude and its signed local components are then

$$b_{3\mathrm{D}} = \operatorname{atan2}(t, \hat{m}_z), \tag{A.6}$$

$$b_\theta = b_{3\mathrm{D}}\frac{\hat{m}_x}{t}, \qquad b_\phi = b_{3\mathrm{D}}\frac{\hat{m}_y}{t}, \tag{A.7}$$

with both components set to zero when $t = 0$. Here b_ϕ is the physical local azimuthal-arc component, not the raw coordinate residual $\Delta\phi$. The relation $\sqrt{b_\theta^2 + b_\phi^2} = b_{3\mathrm{D}}$ follows directly from the definition. The magnitude measures the displacement of the aligned mean direction and must not be confused with the mean or median of the event-wise opening angles $\Delta\psi$.

At $\theta = 0°$ or $180°$, the magnitude $b_{3\mathrm{D}}$ remains well defined. The orientation of b_θ and b_ϕ, however, depends on the azimuth and roll convention assigned at the pole. Pole-centred bins are marked separately in the figures, and their local components are not given an intrinsic physical azimuthal interpretation.

A.2 Spatial diagnostics

The fitted infinite track has a direction and a support point, but no physical neutrino-interaction vertex. Let $\Delta\mathbf{r} = \mathbf{r}_{\mathrm{reco}} - \mathbf{r}_{\mathrm{true}}$. The Euclidean support-point distance and its decomposition relative to the true direction are

$$d_{\mathrm{support}} = \|\Delta\mathbf{r}\|, \tag{A.8}$$

$$d_{\mathrm{support},\parallel} = \Delta\mathbf{r}\cdot\hat{\mathbf{u}}_{\mathrm{true}}, \tag{A.9}$$

$$d_{\mathrm{support},\perp} = \left\|\Delta\mathbf{r} - d_{\mathrm{support},\parallel}\hat{\mathbf{u}}_{\mathrm{true}}\right\|. \tag{A.10}$$

These quantities describe the chosen reference points, not physical vertices. In particular, a displacement along an infinite line can change d_{support} without changing the trajectory itself. They are retained only for exploratory checks in which the support-point convention is relevant.

The minimum distance between the true and reconstructed lines is the preferred spatial diagnostic. For non-parallel directions it is

$$d_{\mathrm{line}} = \frac{|(\mathbf{r}_{\mathrm{reco}} - \mathbf{r}_{\mathrm{true}})\cdot(\hat{\mathbf{u}}_{\mathrm{true}}\times\hat{\mathbf{u}}_{\mathrm{reco}})|}{\|\hat{\mathbf{u}}_{\mathrm{true}}\times\hat{\mathbf{u}}_{\mathrm{reco}}\|}, \tag{A.11}$$

where $\Delta\mathbf{r} = \mathbf{r}_{\mathrm{reco}} - \mathbf{r}_{\mathrm{true}}$. For parallel or anti-parallel directions, the cross-product denominator vanishes and the point-to-line limit is used,

$$d_{\mathrm{line}}^{\parallel} = \|\Delta\mathbf{r} - (\Delta\mathbf{r} \cdot \hat{\mathbf{u}}_{\mathrm{true}})\,\hat{\mathbf{u}}_{\mathrm{true}}\| = \|\Delta\mathbf{r} \times \hat{\mathbf{u}}_{\mathrm{true}}\|\,. \tag{A.12}$$

This quantity is unchanged when either support point is moved along its own line. It is nevertheless complementary to the angular error: two intersecting lines can have zero line distance while retaining different directions.

A.3 Detected information and event survival

The sampled charge Q_i is an integer Poisson realisation. A DOM is counted as hit only when its sampled charge is strictly positive,

$$\mathcal{H} = \{i : Q_i > 0\} = \{i : Q_i \geq 1\}, \qquad N_{\mathrm{hit}} = |\mathcal{H}|, \qquad Q_{\mathrm{tot}} = \sum_{i \in \mathcal{H}} Q_i. \tag{A.13}$$

A DOM with a non-zero expected response but a zero Poisson draw is not part of the event passed to the timing likelihood. Total charge measures the sampled amount of light, while the hit-DOM count also captures how widely that information is distributed over the detector. They are not reconstruction errors, but they are essential when a degradation coincides with a loss of available information.

Event fractions keep this loss visible. Reconstruction is attempted when $N_{\mathrm{hit}} > 0$, so the detected population is defined exactly by

$$N_{\mathrm{det}} = N(N_{\mathrm{hit}} > 0)\,. \tag{A.14}$$

The reconstructed population N_{reco} contains the detected events for which the minimisation returns storable fitted-track parameters. This status confirms that a fit result is available; it does not prove that the global minimum has been found. If N_{gen} events are requested, then

$$f_{\mathrm{det}} = \frac{N_{\mathrm{det}}}{N_{\mathrm{gen}}}, \qquad f_{\mathrm{reco}} = \frac{N_{\mathrm{reco}}}{N_{\mathrm{gen}}}, \qquad f_{\mathrm{reco|det}} = \frac{N_{\mathrm{reco}}}{N_{\mathrm{det}}}. \tag{A.15}$$

Directional and spatial summaries use successful reconstructions, whereas the fractions retain the stated population in their denominator. This distinction is especially important near and beyond the detector boundary.

A.4 Distribution summaries and uncertainty

The median and the 68th and 95th percentiles describe different parts of the non-negative angular-error distribution. Here p_{50} represents typical performance, p_{68} a broader containment scale, and p_{95} the upper tail. For signed coordinate residuals, the mean retains the direction of a shift. For the aligned metric, the normalised mean direction gives $b_{3\mathrm{D}}$ and its two local components. Neither aggregate must be confused with a positive containment percentile.

Uncertainty on a bin-level directional bias is evaluated with a non-parametric bootstrap. Complete events are resampled with replacement and the statistic is recomputed for every replicate,

$$T^{*(b)} = T\left(x_{I_1^{(b)}}, \ldots, x_{I_N^{(b)}}\right), \qquad b = 1, \ldots, B. \tag{A.16}$$

This event-level construction is used for the detector-wide estimates. For the local DOM coordinate-residual summaries, the statistic is first evaluated for each DOM and the resampled units are the per-DOM means. Both the primary 24-DOM curves and the later local-population comparisons use 2,000 DOM-level resamples. The 95% interval is formed from the 2.5th and 97.5th percentiles of the replicate distribution. For the aligned metric, each replicate resamples the vectors $\mathbf{v}_i$ and recomputes their mean direction. The interval measures the sampling uncertainty of the aggregate, not the event-to-event spread. Since $b_{3\mathrm{D}}$ is non-negative by definition, its percentile interval is not used as a formal test of zero bias; the signed components retain the direction of the observed displacement. The angular-error percentiles are reported separately as properties of the reconstructed-event distribution.

Table A.1 gathers the quantities used throughout the analysis and states the role of each one.

Quantity	**Role**	**Interpretation**
$\Delta\psi$; p_{50}, p_{68}, p_{95}	Directional resolution and tail	Non-negative event-distribution summaries.
$b_{3\mathrm{D}}$; b_θ, b_ϕ	Aligned directional bias	Magnitude and signed local components of the aligned mean direction; bootstrap intervals quantify uncertainty on the aggregate.
$\Delta\theta$, $\Delta\phi$	Coordinate-residual diagnostics	Reconstructed minus true; useful only with the stated coordinate convention.
d_{line}	Spatial agreement of infinite tracks	Read together with angular error; used when the geometry requires it.
d_{support}, $d_{\mathrm{support},\perp}$	Exploratory support-point diagnostics	Describe the chosen track reference points, not a physical vertex.
Q_{tot}, N_{hit}	Available optical information	Sampled light amount and sensor multiplicity supplied to the fit.
f_{det}, f_{reco}	Event survival	Prevents good surviving events from hiding detection or fit losses.

Table A.1: Metrics retained for interpreting the reconstruction results.

Appendix B

Complete overview dashboards

The figures gathered in this appendix provide a complete view of the final production scans. They complement the results presented in Chapter 6 without repeating their detailed description. The global set covers the depth (Figure B.1), azimuth (Figure B.2), zenith (Figure B.3), detector-centred radial (Figure B.4), local radial (Figure B.5) and local vertical (Figure B.6) scans. The directional-bias set covers azimuth (Figure B.7), zenith (Figure B.8), depth (Figure B.9), detector-centred radial (Figure B.10), local radial (Figure B.11) and local vertical (Figure B.12).

B.1 Global reconstruction dashboards

The first set of dashboards brings together the angular error, minimum line-to-line distance, detection and reconstruction fractions, number of hit DOMs and total detected charge. This makes it possible to compare the reconstruction quality directly with the amount of optical information available in each scan.

The two local dashboards use the extended 42-DOM diagnostic introduced in Section 7.7; the primary local results in Section 6.6 use the separate 24-DOM production.

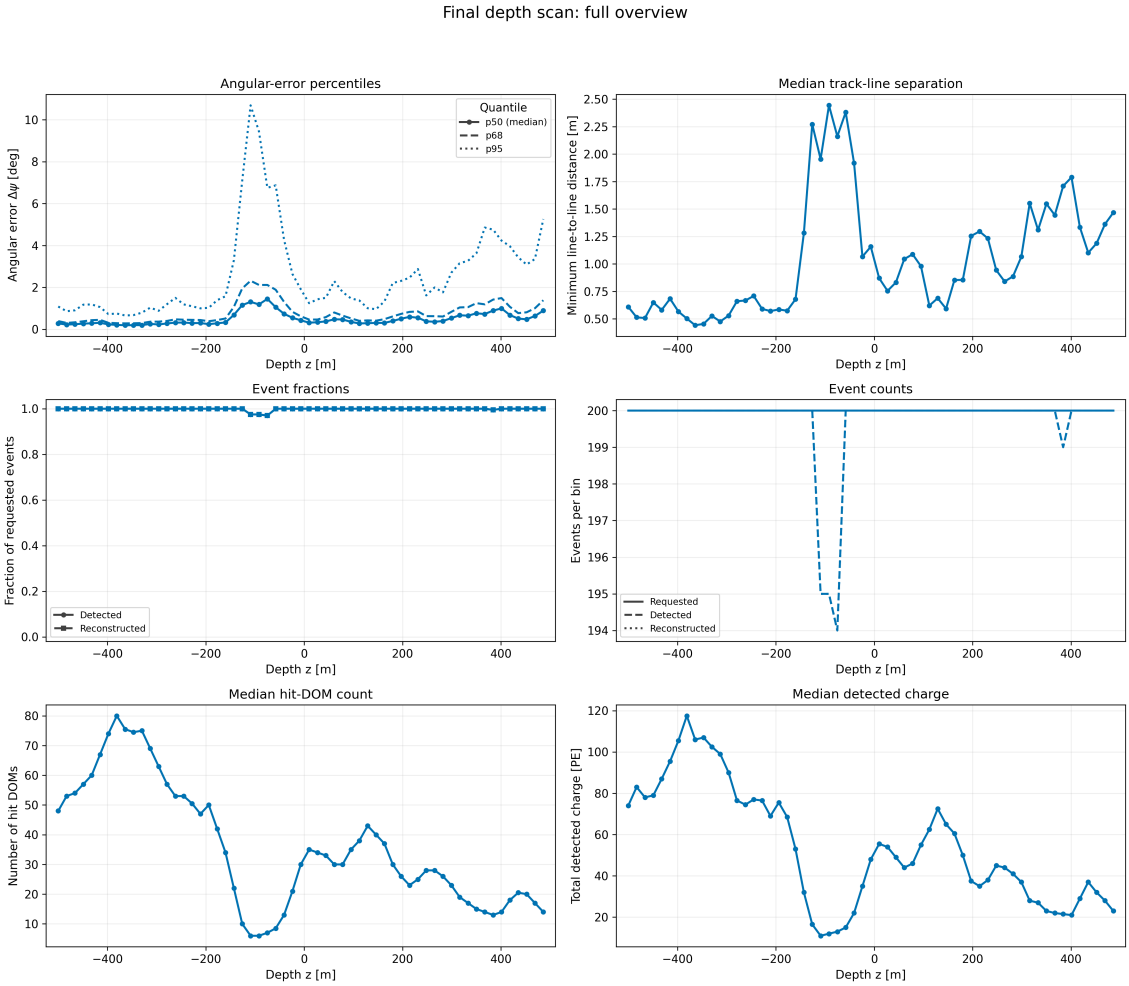


Figure B.1: Complete depth-scan dashboard.

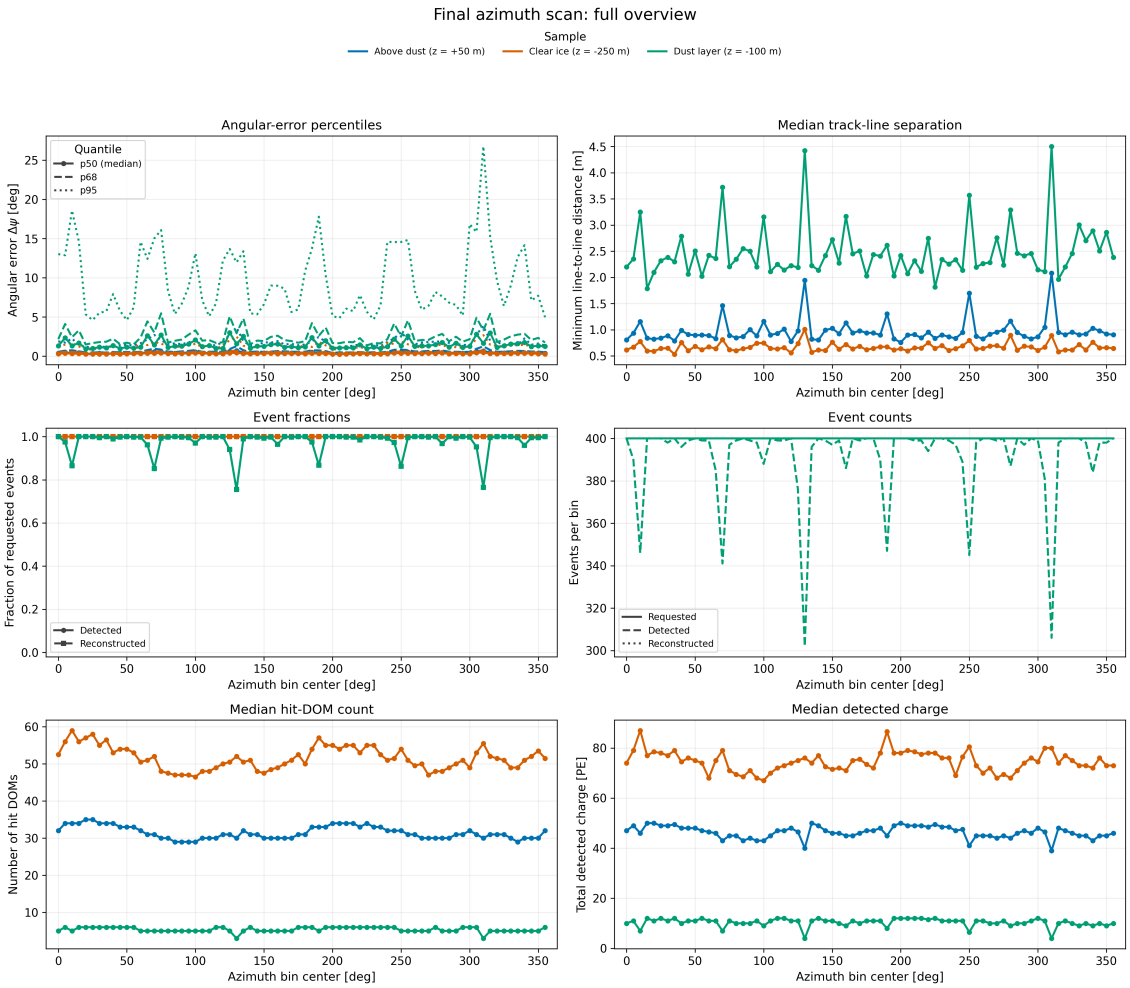


Figure B.2: Complete azimuth-scan dashboard.

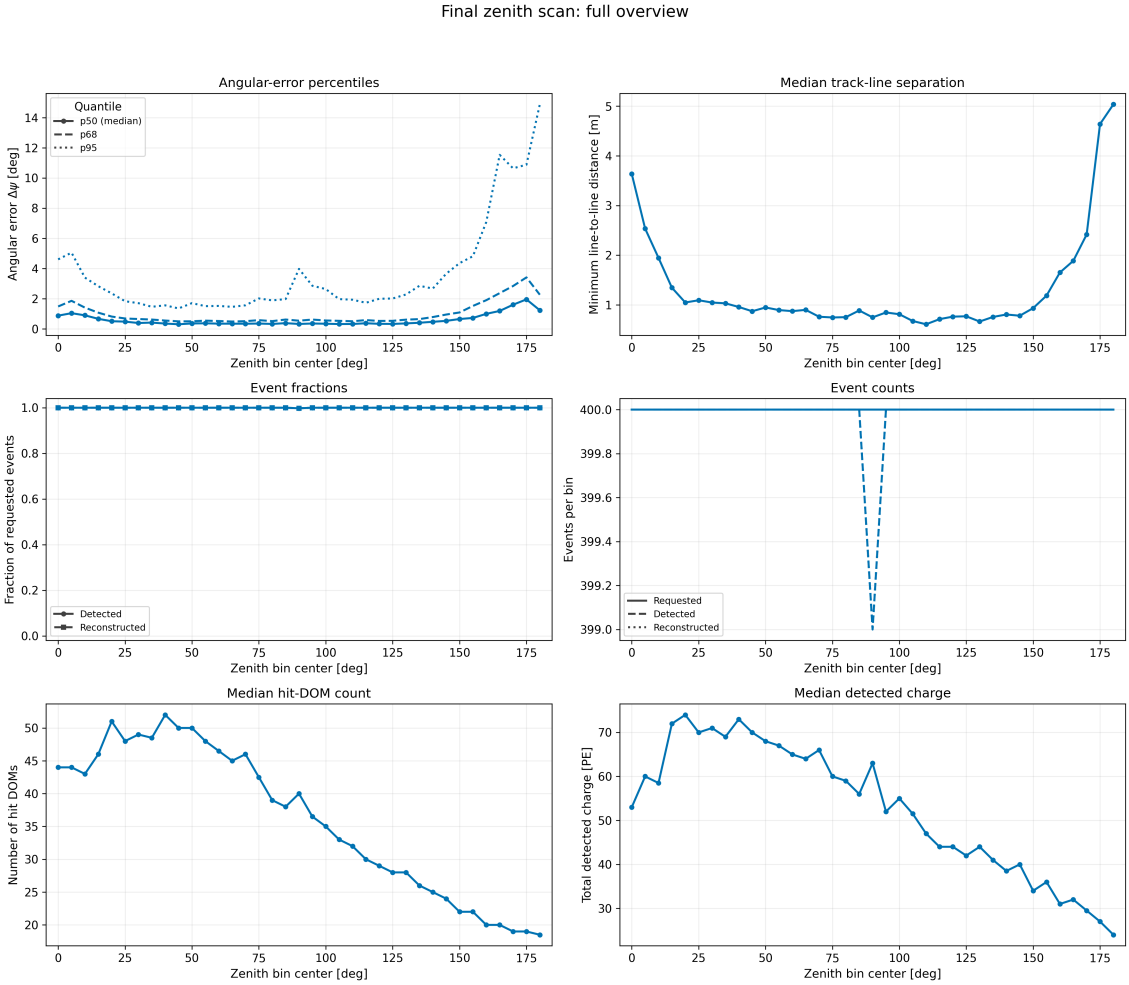


Figure B.3: Complete zenith-scan dashboard.

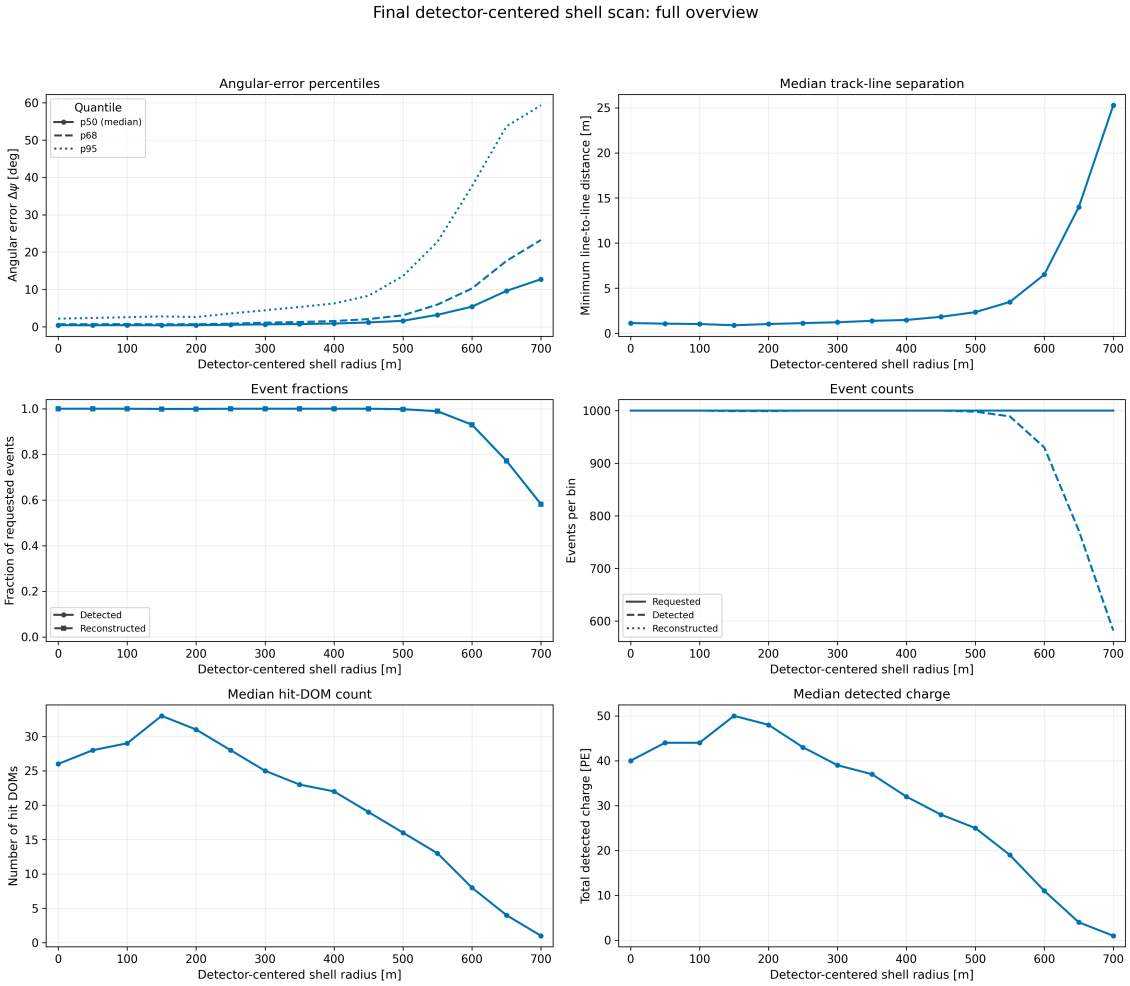


Figure B.4: Complete detector-centred radial-scan dashboard.

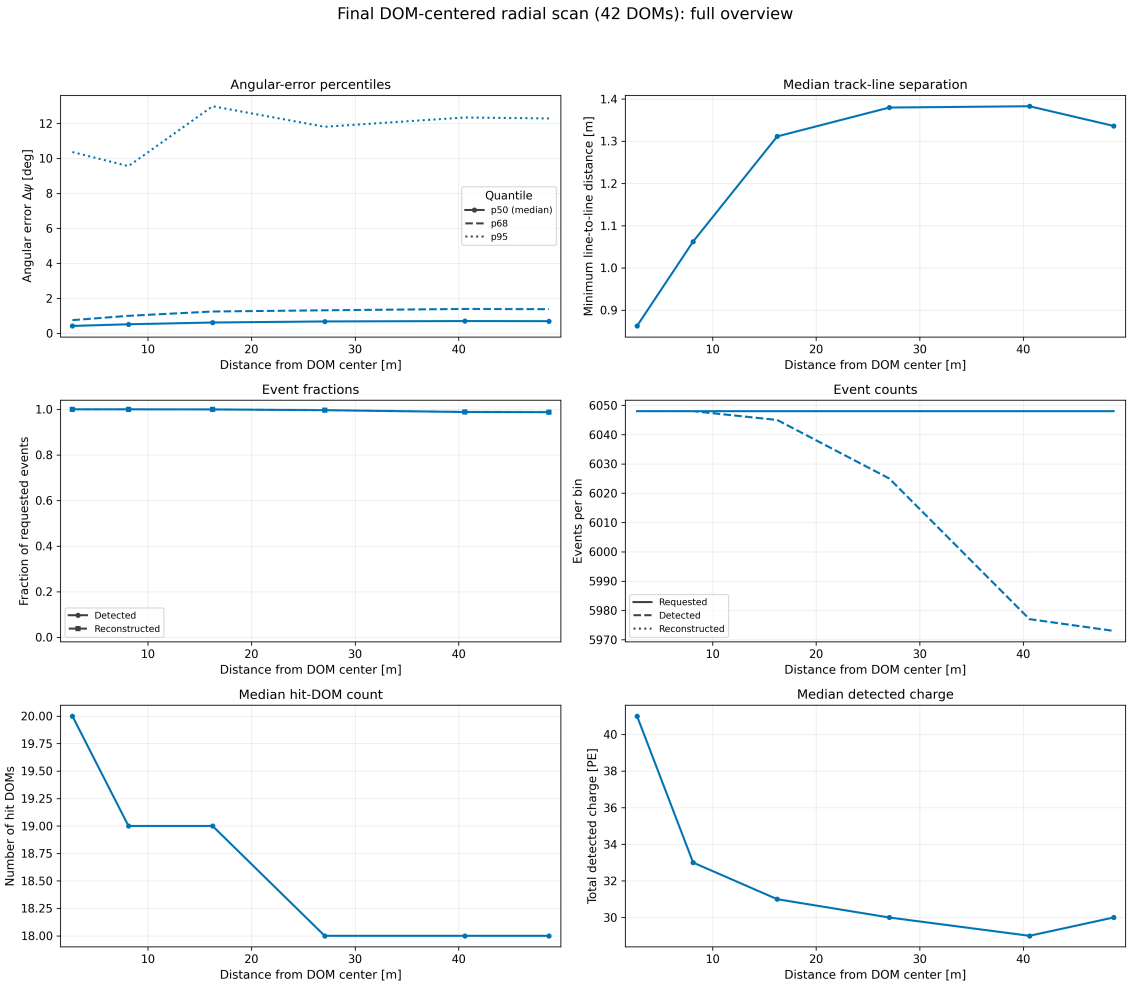


Figure B.5: Complete local radial-scan dashboard for the extended 42-DOM population.

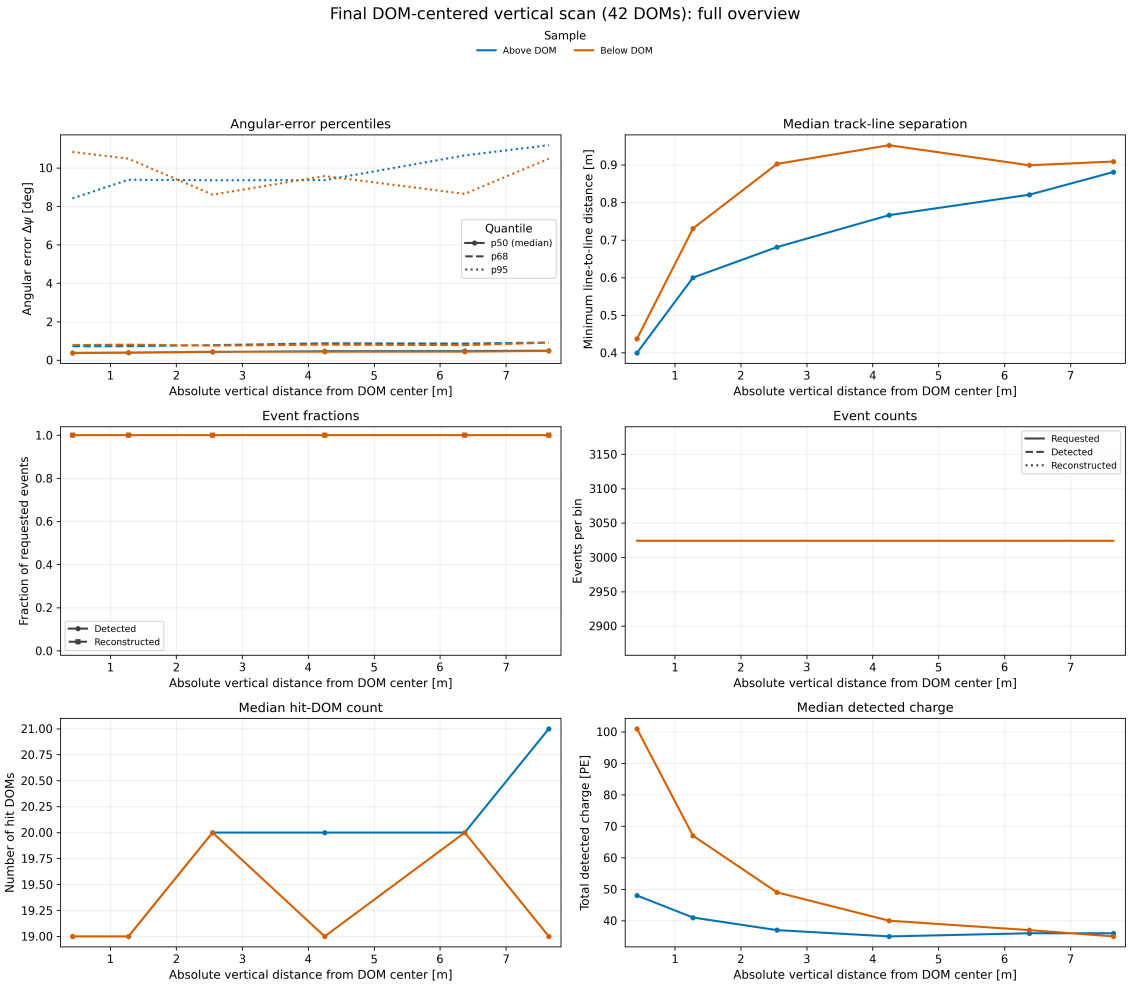


Figure B.6: Complete local vertical-scan dashboard for the extended 42-DOM population.

B.2 Directional-bias dashboards

The second set of dashboards reports directional shifts with their bootstrap uncertainty intervals. For the azimuth, zenith and detector-centred radial scans, the former separate coordinate means are replaced by the aligned three-dimensional magnitude b_{3D} and its roll-fixed local components b_θ and b_ϕ. The depth and local DOM dashboards retain their signed coordinate residuals as diagnostics for the corresponding nearly horizontal configurations. These coordinate residuals should not be identified directly with the aligned components.

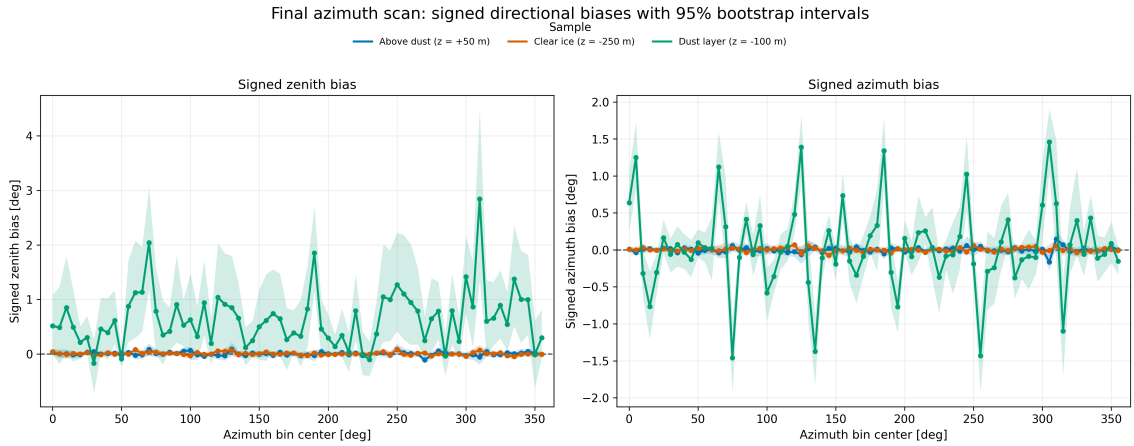


Figure B.7: Aligned three-dimensional directional-bias magnitude and roll-fixed local components for the productions above the dust layer, inside the dust layer and in clear ice. Shaded bands are 95% bootstrap intervals; the three panels use independent vertical scales.

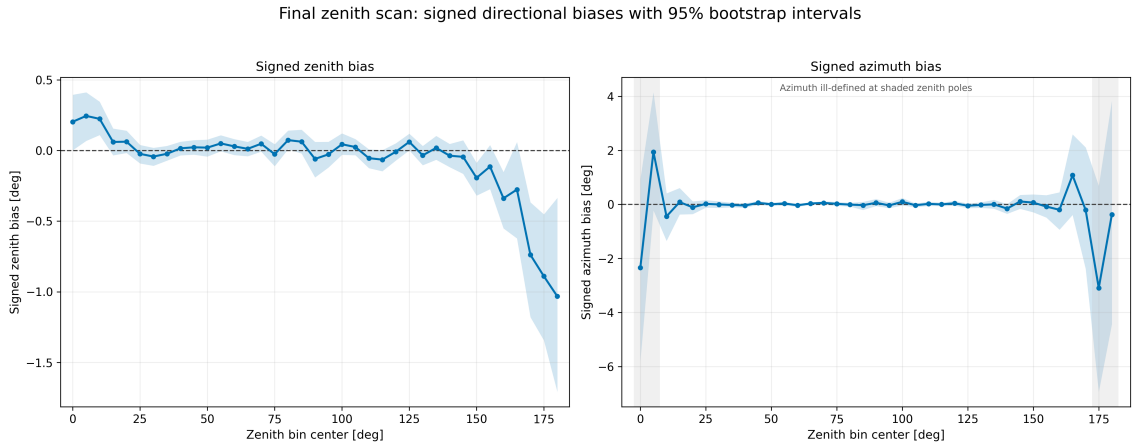


Figure B.8: Aligned three-dimensional directional-bias magnitude and roll-fixed local components for the three vertical sampling centres. Shaded bands are 95% bootstrap intervals and all panels use the same vertical scale. Open endpoint markers identify pole-centred bins, for which the local component orientation is convention-dependent.

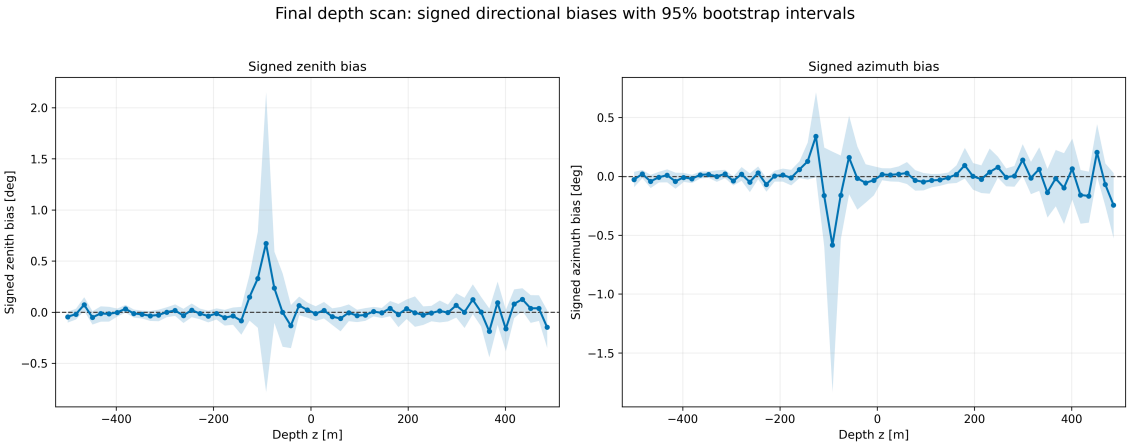


Figure B.9: Signed coordinate residuals in the depth scan.

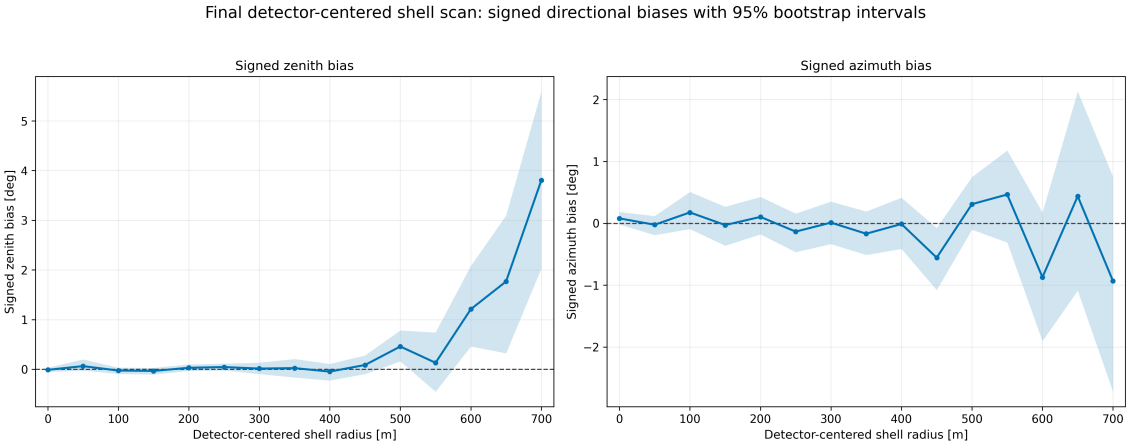


Figure B.10: Aligned three-dimensional directional-bias magnitude and roll-fixed local components in the detector-centred radial scan. Shaded bands are 95% bootstrap intervals.

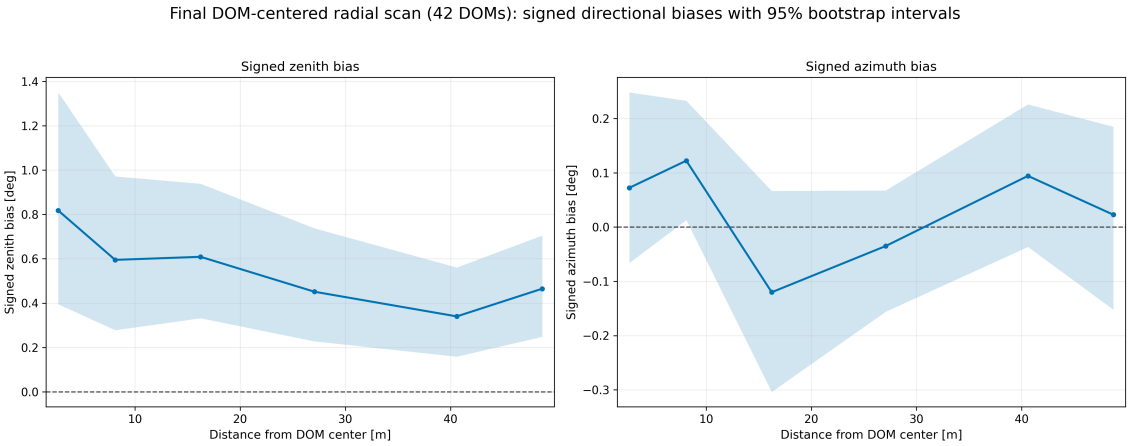


Figure B.11: Signed coordinate residuals in the local radial scan for the extended 42-DOM population.

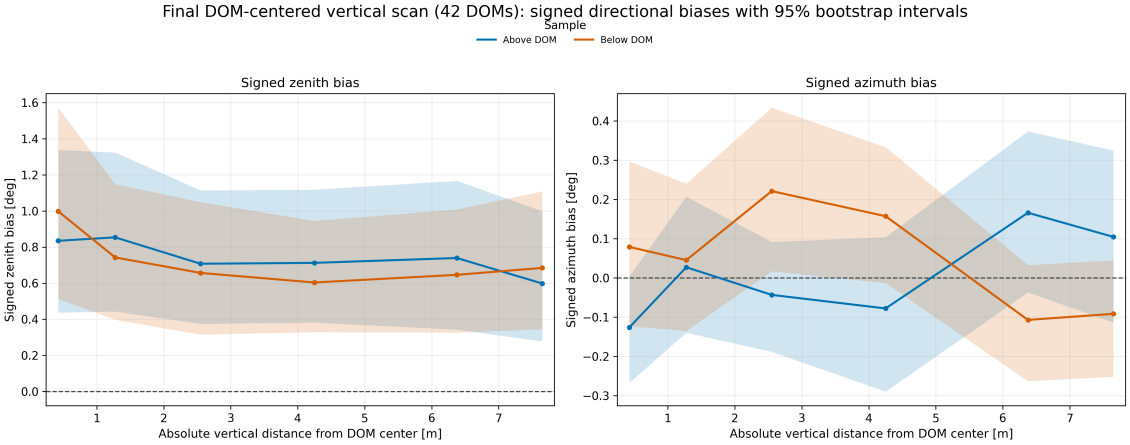


Figure B.12: Signed coordinate residuals in the local vertical scan for the extended 42-DOM population.

Appendix C

Computational Workflow, Reconstruction Blocks and Execution Details

This appendix describes the indispensable computational blocks used by the reconstruction and how they are combined in every scan. The internal file and routine names are not needed to understand the calculation: the same physical sequence is used throughout, while only the sampling of the tested coordinate changes.

C.1 Common event-level workflow

Figure C.1 summarises one production sample. A scan first draws a true track, then produces a synthetic event, fits a candidate track, compares it with the retained truth and stores the event-level result. Bin-wise statistics and figures are calculated only after this sequence has been completed.

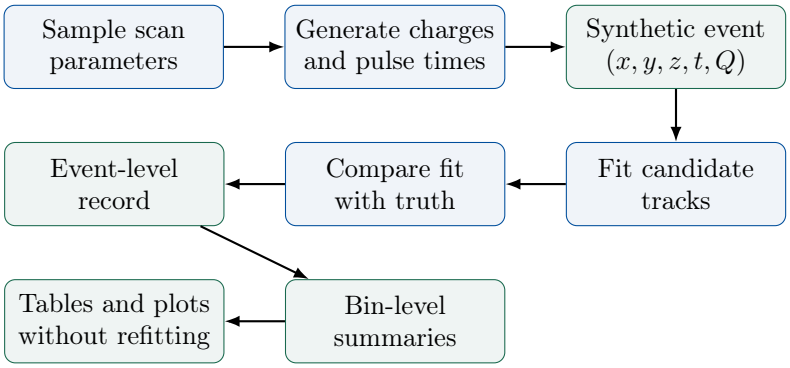


Figure C.1: Common computational sequence for one production sample.

C.2 Indispensable reconstruction blocks

Detector and model initialisation. The detector geometry, the trained charge and timing responses, and the numerical fitter are loaded once before the event loop.

This ensures that every event uses the same DOM positions and model parameters and avoids rebuilding the reconstruction for each stochastic realisation.

Charge evaluation. For a true track, the charge-response block evaluates the expected unscaled charge at every DOM from its geometry relative to the track. In the implemented generation convention, this response is retained only when the perpendicular distance from the DOM to the infinite track satisfies $d_{\perp,i} < 400$ m; it is set to zero for $d_{\perp,i} \geq 400$ m. The fixed light-yield scale is then applied and an integer charge is drawn from a Poisson distribution. A DOM belongs to the hit set only if the sampled charge satisfies $Q_i > 0$, equivalently $Q_i \geq 1$. The resulting hit and detected-event populations are defined in Section A.3. This block determines both the amount of detected light and which sensors are available to the timing fit.

Pulse-time generation. For each hit DOM, the track–DOM geometry gives the direct Cherenkov arrival time. The timing response supplies the delay distribution, from which the first arrival is sampled using the DOM charge as the effective multiplicity. A 3 ns Gaussian timing fluctuation is added. The synthetic event then contains one position, first-pulse time and sampled charge for every hit DOM.

DOM–track timing evaluation. During reconstruction, the same geometrical quantities are recomputed for each trial track. The trained timing response converts them into the parameters of the analytical delay distribution. This block is evaluated repeatedly as the candidate direction and support point move during minimisation.

MPE likelihood. The observed charge sets the MPE multiplicity of each hit DOM and conditions the first-arrival density. The multiplicity used in the likelihood is

$$n_i = \mathrm{clip}(Q_i, 1, 10000). \tag{C.1}$$

A small uniform component protects the objective against vanishing timing probabilities,

$$\widetilde{p}_i = 0.99\, p_i^{\mathrm{MPE}} + 0.01 \left(\frac{1}{6000\ \mathrm{ns}} \right), \qquad \mathcal{J}(\Theta) = -2 \sum_{i \in \mathcal{H}} \ln \widetilde{p}_i. \tag{C.2}$$

The charge is used to shape the timing term; no separate charge-likelihood term is fitted. The objective value is interpreted within this fixed timing convention and not compared directly between events with very different hit multiplicities.

Fitter. The fitted variables are the zenith, azimuth and three coordinates of the track support point. Seven nearby support-point initialisations are tested. For each one, a discrete reference-time prescan is followed by a BFGS minimisation, and the solution with the smallest objective is retained. These initialisations are several attempts to reconstruct the same event, not additional generated events.

Event evaluation. After the fit, the reconstructed track is compared with the retained true track using the metrics in Appendix A. The event record keeps the true and reconstructed parameters, detected charge, hit-DOM multiplicity, reconstruction status and scan coordinate before any aggregation. This is what allows a new percentile, signed coordinate residual, aligned bias or correlation to be examined without generating and fitting the event again.

C.3 How the scans reuse the workflow

The azimuth, zenith, depth, detector-centred radial and local-DOM studies all use the sequence in Figure C.1. Their only substantive difference is the first block: each study controls one coordinate and samples its nuisance variables according to the protocol stated in Chapter 5. Once a track has been drawn, charge generation, pulse-time sampling, likelihood evaluation, fitting and event-level storage proceed in the same way.

Large productions can be divided into consecutive subsets and recombined at event level. This was required when retaining an entire campaign and its intermediate objects in one process exceeded the available memory. Partitioning changes memory use and failure recovery, but it does not create new physical regions or alter the quantity being measured, provided that the sampling and reconstruction settings remain identical.

C.4 Execution constraints and retained records

The analysis was run under WSL Ubuntu using JAX for the neural-network evaluation, automatic differentiation and likelihood minimisation [38]. NumPy, pandas, SciPy and matplotlib were used for event generation, data processing, statistical analysis and plotting.

The simulations were performed on an NVIDIA GeForce RTX 3060 Laptop GPU with 6 GB of memory, an AMD Ryzen 7 5800H CPU and 16 GB of DDR4-3200 RAM. The available memory and the way JAX handles its inputs influenced how the production scripts were organised, without changing the physical definition of the scans.

JAX input shapes. Each synthetic event contains only the DOMs with a sampled charge $Q_i > 0$. The number of rows given to the reconstruction depends on the number of hit DOMs:

$$\text{input shape} = (N_{\text{hit}}, N_{\text{features}}) \,. \tag{C.3}$$

For example, two events treated with the same reconstruction can have input shapes such as $(276, 5)$ and $(285, 5)$. Since JAX compiles its operations for compatible input shapes, encountering a new shape can require an additional compilation. The first call is then slower, while later calls with a compatible shape can reuse the compiled operation.

Figure C.2 illustrates why two events reconstructed with the same method can nevertheless trigger different JAX compilations.

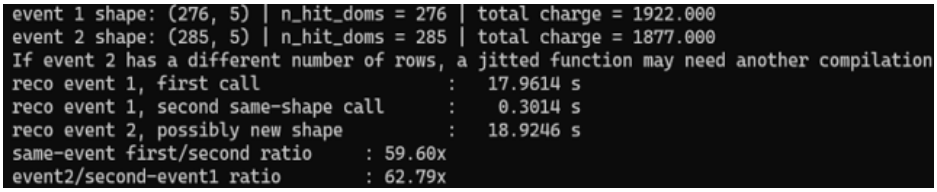


Figure C.2: Example of the shape dependence of JAX compilation. The number of hit DOMs changes the first dimension of the event array. A first call can include compilation time, whereas a compatible later call can reuse the compiled operation.

This is why the detector model and fitter were initialised before the event loop and then reused. The same charge, timing, likelihood and fitting blocks were also kept for all scans. Their structure remained unchanged; only the parameters sampled at the beginning of the workflow depended on the scan being performed.

The variable number of hit DOMs also affects the computational cost. An event with a different input shape can require a new compilation, and an event with more hit DOMs requires more values to be processed by the likelihood. The total execution time is consequently not strictly proportional to the number of generated events. This is only a computational effect: it does not change the likelihood or the definition of the reconstructed quantities.

Some large productions were divided into smaller consecutive subsets because keeping the complete sample and its intermediate results in memory was not possible with the available RAM. All subsets used the same generation and reconstruction settings and were recombined at event level before calculating the final statistics. This division changed only the execution of the production, not the physical population being studied.

Requirements for reproduction. Reproducing a production requires the same detector geometry, trained-model file, software versions, sampling rules, numerical settings and random seeds. The event-level results were stored separately from the final plots, which allows new percentiles, aligned biases, coordinate residuals or correlations to be calculated without generating and reconstructing the events again. The appendix documents the workflow and the conditions needed for reproduction, but it is not a self-contained software release: the exact version identifiers, model file, production configurations and seed records must accompany the analysis code for a complete rerun.

Appendix D

AI Usage

I acknowledge having used artificial intelligence tools as assistance during the writing process, including ChatGPT.

More specifically, Codex was used to assist with code development, while ChatGPT was used for other supporting tasks.

All research, scientific reasoning, interpretation, and analysis presented in this thesis were carried out by me. I take full responsibility for the scientific content and conclusions of this work.

Bibliography

[1] M. G. Aartsen et al. Evidence for high-energy extraterrestrial neutrinos at the IceCube detector. *Science*, 342(6161):1242856, 2013. doi: 10.1126/science.1242856.

[2] Markus Ahlers and Francis Halzen. Opening a new window onto the universe with IceCube. *Progress in Particle and Nuclear Physics*, 102:73–88, September 2018. ISSN 01466410. doi: 10.1016/j.ppnp.2018.05.001.

[3] J. A. Formaggio and G. P. Zeller. From eV to EeV: Neutrino cross sections across energy scales. *Reviews of Modern Physics*, 84(3):1307–1341, September 2012. ISSN 0034-6861, 1539-0756. doi: 10.1103/RevModPhys.84.1307.

[4] Teresa Montaruli. Selected results from IceCube. *Journal of Physics: Conference Series*, 2429:012026, February 2023. doi: 10.1088/1742-6596/2429/1/012026.

[5] Kai Schatto. *Stacked Searches for High-Energy Neutrinos from Blazars with IceCube*. PhD thesis, Johannes Gutenberg-Universität Mainz, Mainz, Germany, 2014.

[6] Hans Niederhausen. High Energy Track Reconstruction with TriplePandelMPE. Internal presentation, IceCube Collaboration, April 2025.

[7] S. Bilenky. Neutrino oscillations: From a historical perspective to the present status. *Nuclear Physics B*, 908:2–13, July 2016. ISSN 0550-3213. doi: 10.1016/j.nuclphysb.2016.01.025.

[8] C. L. Cowan et al. Detection of the Free Neutrino: A Confirmation. *Science*, 124(3212):103–104, July 1956. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.124.3212.103.

[9] G. Danby et al. Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos. *Physical Review Letters*, 9(1):36–44, July 1962. ISSN 0031-9007. doi: 10.1103/PhysRevLett.9.36.

[10] K. Kodama et al. Observation of tau neutrino interactions. *Physics Letters B*, 504(3):218–224, April 2001. ISSN 03702693. doi: 10.1016/S0370-2693(01)00307-0.

[11] Bruce T. Cleveland et al. Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector. *The Astrophysical Journal*, 496(1):505–526, March 1998. ISSN 0004-637X, 1538-4357. doi: 10.1086/305343.

[12] A. B. McDonald. Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in SNO. In *AIP Conference Proceedings*, volume 646, pages 43–58, Waterloo, Ontario (Canada), 2002. AIP. doi: 10.1063/1.1524553.

[13] Adam Burrows and James M. Lattimer. Neutrinos from SN 1987A. *The Astrophysical Journal*, 318:L63, July 1987. ISSN 0004-637X, 1538-4357. doi: 10.1086/184938.

[14] Georg G. Raffelt, Hans-Thomas Janka, and Damiano F. G. Fiorillo. Neutrinos from core-collapse supernovae. arXiv:2509.16306, 2025.

[15] Cush. Standard Model of Elementary Particles. `https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`, September 2019. Accessed 24 July 2026.

[16] Mary Hall Reno. High-Energy to Ultrahigh-Energy Neutrino Interactions. *Annual Review of Nuclear and Particle Science*, 73(1):181–204, September 2023. ISSN 0163-8998, 1545-4134. doi: 10.1146/annurev-nucl-111422-040200.

[17] B. Abi et al. Prospects for beyond the Standard Model physics searches at the Deep Underground Neutrino Experiment: DUNE Collaboration. *The European Physical Journal C*, 81(4):322, April 2021. ISSN 1434-6044, 1434-6052. doi: 10.1140/epjc/s10052-021-09007-w.

[18] IceCube Collaboration et al. Characterization of the astrophysical diffuse neutrino flux using starting track events in IceCube. *Physical Review D*, 110(2): 022001, July 2024. doi: 10.1103/PhysRevD.110.022001.

[19] Federica Bradascio and Thorsten Glüsenkamp. An improved muon track reconstruction for IceCube. In *Proceedings of the 36th International Cosmic Ray Conference (ICRC2019)*, page 846. Sissa Medialab, July 2021. doi: 10.22323/1.358.0846.

[20] J. Ahrens et al. Muon track reconstruction and data selection techniques in AMANDA. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 524(1-3): 169–194, May 2004. ISSN 01689002. doi: 10.1016/j.nima.2004.01.065.

[21] Christian Spiering. Neutrino Detectors Under Water and Ice. In Christian W. Fabjan and Herwig Schopper, editors, *Particle Physics Reference Library*, pages 785–822. Springer International Publishing, Cham, 2020. ISBN 978-3-030-35317-9 978-3-030-35318-6. doi: 10.1007/978-3-030-35318-6_17.

[22] R. Abbasi et al. The design and performance of IceCube DeepCore. *Astroparticle Physics*, 35(10):615–624, May 2012. ISSN 09276505. doi: 10.1016/j.astropartphys.2012.01.004.

[23] IceCube Collaboration et al. The IceCube Neutrino Observatory: Instrumentation and Online Systems. *Journal of Instrumentation*, 12(03):P03012–P03012, March 2017. ISSN 1748-0221. doi: 10.1088/1748-0221/12/03/P03012.

[24] William Luszczak. TXS 0506+056 with Updated IceCube Data. In *Proceedings of the 38th International Cosmic Ray Conference (ICRC2023)*, page 1465. Sissa Medialab, 2023. doi: 10.22323/1.444.1465.

[25] IceCube Collaboration et al. Evidence for neutrino emission from the nearby active galaxy NGC 1068. *Science*, 378(6619):538–543, November 2022. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.abg3395.

[26] IceCube Collaboration et al. Observation of high-energy neutrinos from the Galactic plane. *Science*, 380(6652):1338–1343, June 2023. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.adc9818.

[27] Andrii Terliuk and IceCube Collaboration. Atmospheric neutrino oscillations with IceCube. In *Proceedings of Neutrino Oscillation Workshop — PoS(NOW2018)*, page 007, Rosa Marina (Ostuni, Brindisi, Italy), April 2019. Sissa Medialab. doi: 10.22323/1.337.0007.

[28] IceCube Collaboration et al. Detection of a particle shower at the Glashow resonance with IceCube. *Nature*, 591:220–224, March 2021. doi: 10.1038/s41586-021-03256-1.

[29] IceCube Collaboration. The IceCube Neutrino Observatory: Detector, n.d. URL `https://icecube.wisc.edu/science/icecube/`. Accessed 4 August 2026.

[30] IceCube Collaboration. IceCube-86 (78+8) Interstring (Surface) Distances, 2011. URL `https://wiki.icecube.wisc.edu/index.php/File:Distances.i86.jpg`. Accessed 4 August 2026.

[31] Tianlu Yuan. Detecting neutrinos in IceCube with Cherenkov light in the South Pole ice. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1054:168440, September 2023. ISSN 0168-9002. doi: 10.1016/j.nima.2023.168440.

[32] The IceCube Collaboration et al. Calibration and Characterization of the IceCube Photomultiplier Tube. `https://arxiv.org/abs/1002.2442v1`, February 2010.

[33] R. Abbasi et al. The IceCube data acquisition system: Signal capture, digitization, and timestamping. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 601(3):294–316, April 2009. ISSN 01689002. doi: 10.1016/j.nima.2009.01.001.

[34] Hamamatsu Photonics. R7081 Photomultiplier Tube, n.d. URL `https://www.hamamatsu.com/eu/en/product/optical-sensors/pmt/pmt_tube-alone/head-on-type/R7081.html`. Official product page; accessed 14 August 2026.

[35] Joschua Hellemeier. Effect of noise on IceCube's muon track reconstruction. Bachelor's thesis, Johannes Gutenberg-Universität Mainz, Mainz, Germany, November 2011.

[36] M.G. Aartsen et al. Measurement of South Pole ice transparency with the IceCube LED calibration system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 711:73–89, May 2013. ISSN 01689002. doi: 10.1016/j.nima.2013.01.054.

[37] Dirk Pandel. Bestimmung von wasser- und detektorparametern und rekonstruktion von myonen bis 100 TeV mit dem baikal-neutrinoteleskop NT-72. Master's thesis, Humboldt-Universität zu Berlin, Berlin, Germany, 1996.

[38] JAX authors. JAX: Composable Transformations of Python+NumPy Programs, 2018. URL `https://github.com/jax-ml/jax`. Accessed 7 July 2026.

[39] Donald E. Groom, Nikolai V. Mokhov, and Sergei I. Striganov. MUON STOPPING POWER AND RANGE TABLES 10 MeV–100 TeV. *Atomic Data and Nuclear Data Tables*, 78(2):183–356, July 2001. ISSN 0092640X. doi: 10.1006/adnd.2001.0861.

[40] Leif Rädel and Christopher Wiebusch. Calculation of the Cherenkov light yield from low energetic secondary particles accompanying high-energy muons in ice and water with Geant 4 simulations. *Astroparticle Physics*, 38:53–67, October 2012. ISSN 09276505. doi: 10.1016/j.astropartphys.2012.09.008.

[41] B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), January 1979. ISSN 0090-5364. doi: 10.1214/aos/1176344552.

[42] Matteo Agostini and Birgit Neumair. Statistical methods applied to the search of sterile neutrinos. *The European Physical Journal C*, 80(8):750, August 2020. ISSN 1434-6044, 1434-6052. doi: 10.1140/epjc/s10052-020-8279-x.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Faculté des sciences

Place des Sciences, 2 bte L6.06.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/sc**