

Markov-Chain Monte-Carlo Reconstruction for cascade-like events in IceCube

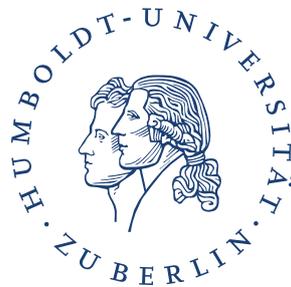
Diplomarbeit

zur Erlangung des akademischen Grades
Diplom-Physiker
(Dipl.-Phys.)

eingereicht an der
Humboldt-Universität zu Berlin
Mathematisch-Naturwissenschaftliche Fakultät I

von

Jakob van Santen
geboren am 27. September 1983 in Madison, WI, USA



1. Gutachter: Prof. Dr. Marek Kowalski
2. Gutachter: Prof. Dr. Hermann Kolanoski

Berlin, den 22. Januar 2010

Contents

1	Introduction	3
1.1	Motivation	3
1.2	Overview	3
1.3	Neutrinos	4
1.4	Motivation for neutrino astronomy	4
1.5	IceCube	5
2	Detecting neutrinos with IceCube	7
2.1	Deep inelastic neutrino-nucleon scattering	7
2.2	Neutrino-electron scattering	8
2.3	Neutrino event signatures	8
2.4	The Cherenkov Effect	10
2.5	Physics of showers	10
2.5.1	Electromagnetic showers	11
2.5.2	Hadronic showers	11
2.6	Light propagation in Antarctic ice	12
2.7	Light detection and readout	12
2.7.1	Photomultiplier tubes	12
2.7.2	The Digital Optical Module	13
3	The IceCube cascade simulation chain	16
3.1	The IceTray software framework	16
3.2	Event generation	16
3.3	Light propagation: Photonics	18
3.4	Hit generation	20
3.5	PMT simulation	20
3.6	DOM simulation	21
4	Cascade reconstruction	22
4.1	Event parameters	22
4.1.1	Coordinate system	22
4.1.2	Timing	22
4.1.3	Visible energy	22
4.2	Low-level reconstruction	23
4.2.1	DOM calibration	23
4.2.2	Pulse reconstruction	24
4.3	First-guess algorithms and filtering	26
5	Likelihood methods	27
5.1	Statistical inference	27
5.2	Bayes' Theorem	27
5.3	Maximum likelihood estimators	28
5.4	A likelihood function for cascades	28
5.5	Error estimates from the likelihood function	29
5.6	Modelling the detector response	30
5.7	Calculating the error volume	31
6	Markov Chains	33

6.1	Theory	33
6.2	Implementation	34
6.2.1	The Metropolis Algorithm	34
6.2.2	Python implementation	35
7	Application to cascade reconstruction	39
7.1	Simplified reconstruction problem	39
7.2	Benchmark data set and implementation	39
7.3	Implicit approximations in the likelihood function	43
7.4	Deconstructing reconstruction	46
7.5	Interpolation with B-splines	49
7.6	Simplified simulation and reconstruction with spline interpolation	51
7.6.1	Practical use of spline tables	51
7.6.2	Baseline reconstruction	53
7.6.3	Total amplitude correction	54
7.6.4	Zero-charge bins	54
7.7	Reconstruction with electronics simulation	55
7.7.1	Correcting for the PMT charge response	55
7.7.2	Reconstruction in restricted dimensions	58
7.7.3	Reconstruction in all dimensions	58
8	Conclusion	64
A	Total event charge as a function of energy	70
B	Further resolution plots	71

1 Introduction

1.1 Motivation

This work deals with the problem of reconstructing the location, direction, and energy of an electromagnetic cascade in a large, mostly transparent, and sparsely-instrumented medium (such as the glacial ice under the South Pole) from the Cherenkov light signatures of the relativistic charged particles in the cascade. The information about an energetic event in the ice is carried by the arrival times of Cherenkov photons at detectors at fixed locations, and is smeared out by scattering caused by dust and bubbles in the ice. As such, the resolution of a reconstruction algorithm can be highly variable, and it can be useful to know the quality of the reconstruction for an individual event.

The problem of extracting information about the underlying event from a small number of arrival times is best approached using the methods of statistical inference. In particular, it is possible to define a statistical model of the measurement process and an associated *likelihood* function that describes the degree to which a given data set supports a particular event hypothesis. This function can be used to describe both the best-fitting hypothesis (the reconstruction) and the range of hypothesis that are compatible with the data. This allows one to calculate the resolution of the reconstruction for individual events.

This work presents a method for estimating the error of a maximum-likelihood estimate and shows how this can be applied to the reconstruction of cascade-like events in the ICECUBE neutrino detector. Furthermore, it shows how this method can be used to find errors in the formulation of the likelihood function used in the reconstruction. Finally, it presents some corrections to the likelihood function that produce better agreement between simulation and reconstruction as well as produce more accurate estimates of the reconstruction error.

1.2 Overview

This section introduces the neutrino, presents a brief motivation for neutrino astronomy, and provides a broad overview of the ICECUBE detector. Section 2 introduces the physical processes that give rise to detectable signals in ICECUBE and provides relevant details of the readout electronics. Section 3 discusses the software used to simulate cascade events, as it is relevant to the special simulation chain that will be presented later. Section 4 then introduces the parameters of the cascade model that will be used for reconstruction; it also gives an overview of the low-level reconstructions which are the algorithms used to generate physically meaningful data for more advanced reconstructions from the raw, uncalibrated data stream. Section 5 gives a theoretical motivation for the Method of Maximum Likelihood as a basis for the most powerful reconstruction algorithms as well as a likelihood function for cascade-like events that was developed by [29]. It introduces the concept of error and ways of estimating the error from the shape of the likelihood function for each individual event. The most general way of estimating the error by the latter method is to perform an N -dimensional integral over the entire parameter space. Section 6 introduces the theory of Markov chains and their use in generating a sample of points that can be used to efficiently approximate such an integral. It goes on to introduce the Metropolis Algorithm as well as the relevant details of its implementation in the Python programming language, particularly the criteria for determining correlation and convergence of the chain.

Section 7 details the actual work done for this thesis. First, the Markov-chain reconstruction is applied to data produced by the cascade event simulation with the existing implementation of the cascade likelihood function and is shown to give inconsistent results. After listing the implicit assumptions and approximations in the likelihood function, it describes how they were removed from the simulation and reconstruction one by one until consistency was achieved. Some detector effects were added to the simulation and reconstruction again, with appropriate and well-motivated corrections to the likelihood function. The detector resolution is shown to be significantly better in this simplified, corrected model, with reconstruction errors consistent with the estimates from the likelihood function.

1.3 Neutrinos

Neutrinos are the neutral partners of the charged leptons; they exist in three flavors, corresponding to the e , μ , and τ leptons. They are the lightest of the fermions; the current upper bound on the ν_e mass is roughly 2 eV. Since they carry neither electric nor color charge, they participate only in gravitational and weak interactions. This allows them to propagate through the interstellar medium (indeed, most matter) without being deflected or absorbed. Neutrinos have another interesting property: their flavor eigenstates do not match their mass eigenstates. Neutrinos produced in weak interactions in definite flavor states will oscillate to other flavors as they propagate through space [39].

1.4 Motivation for neutrino astronomy

The field of astroparticle physics is largely concerned with exploring the mechanisms by which energy is transferred from larger structures in the universe, such as stars and galactic nuclei, to the highly energetic elementary particles, such as protons, electrons, photons, and neutrinos that are observed from Earth or orbiting satellites. The accelerators producing these particles must be powerful: the Earth is subject to a constant, nearly isotropic flux of charged particles (cosmic rays) with energies up to 10^{20} eV, and point sources of high-energy photons (gamma rays) have been observed with energies up to 10^{14} eV. Even though we know that the accelerators are powerful, their natures remain a mystery: possible mechanisms include shock acceleration [9] in shell-type supernova remnants (SNR) or in the jets of active galactic nuclei (AGN) [26]. Precision observation of high-energy neutrinos can illuminate this mystery by adding information not accessible through either gamma- or cosmic-ray astronomy.

Neutrinos will be produced whenever a beam of high-energy hadrons (like the hadronic cosmic ray flux) impinges on matter or radiation fields. Proton-proton collisions will produce both charged and neutral mesons, the lightest and most common of which are the pions. Neutral pions decay almost immediately ($\tau \sim 10^{-17}$ s [17]) to a pair of photons, while the charged pions, which can only decay weakly, are much longer lived ($\tau \sim 10^{-8}$ s [17]). If the target is dense, then most of the charged pions will interact before they can decay, producing more pions and leading to a hadronic shower in the target. Thin targets are more interesting for astrophysics. When the mean free path in the target is much longer than the decay length $\gamma c\tau$, most of the pions will decay weakly, producing muon- and electron-neutrinos that can escape the target rather than depositing their energy locally. Thin targets convert a portion of a hadronic particle flux into a flux of neutrinos.

Radiation fields also provide a suitable target. Protons may also interact with photons to create a Δ^+ baryon, which can decay to either $p + \pi^0$ or $n + \pi^+$ and thence to photons or

neutrinos as in thin hadronic targets. This group of scenarios is usually referred to as a “cosmic beam dump” in analogy to the conventional method of producing a neutrino beam at a particle accelerator¹.

This makes high-energy neutrino production a hallmark of hadronic accelerators. Beam dumps on gas clouds or thermal photons in the vicinity of a cosmic accelerator will produce both charged and neutral pions. The charged pions decay leptonically, producing neutrinos, and the neutral pions electromagnetically to photons. These photons are an object of principal interest in TeV gamma-ray astronomy (see e.g. [13]). However, they are an ambiguous signature of hadronic cosmic ray acceleration, as they could just as well originate from a leptonic accelerator via inverse Compton scattering. Observation of a high energy neutrino signal then provides an important and unique diagnostic of hadronic acceleration in the source.

Direct observation of cosmic ray particles also presents its difficulties. Unlike gamma rays, the charged cosmic rays are deflected by interstellar and intergalactic magnetic fields and thus do not necessarily point back to their sources. If a cosmic ray source is near a thin target such as a gas cloud, neutrinos would be produced in the beam dump and would reach the Earth undeflected. Since they interact only weakly, the neutrinos will propagate with very little absorption, even through dense intervening clouds, and since they are electrically neutral, they are not deflected by magnetic fields. This makes neutrinos ideal messengers from the universe’s most violent regions.

1.5 IceCube

ICECUBE is cubic-kilometer neutrino detector currently under construction deep in the glacial ice at the South Pole [14]. The instrument itself, when complete, will consist of 5160 Digital Optical Modules (DOMs), glass pressure spheres housing a photomultiplier tube (PMT) and readout electronics. 4800 of these DOMs are arranged on 80 \sim 2.5 km long “strings” spaced on a hexagonal grid 125 m apart as shown in Figure 1. The last kilometer of each string is instrumented with 60 DOMs, spaced 17 m apart. The remaining 360 DOMs make up the DEEPCORE extension at the horizontal center of the detector [38]. The 6 DEEPCORE strings are spaced 72 m apart, instrumented with higher quantum-efficiency DOMs concentrated along the last 350 m of the string, spaced 7 m apart. The DOMs detect the light emitted by charged particles as they move faster than the local speed of light in the ice; this allows the array to detect the sudden appearance of relativistic charged particles in the ice, a signature of a neutrino interaction.

The experiment focuses mainly on the the muons produced in charged-current interactions of muon-neutrinos, which leave long “tracks” of light through the sparsely-instrumented detector due to their large mass and relatively long lifetime. Up-going muon tracks, caused by neutrino interactions, can be distinguished from down-going muons, the vast majority of which are the products of cosmic ray interactions in the atmosphere. In this channel, the detector observes the northern sky through the Earth.

Work is underway to extend the sensitivity and resolution of the detector to other neutrino flavors (e and τ), whose interaction products deposit energy in a relatively small portion of the detector. Since electrons and tau leptons have a short range in ice, this detection channel is also

¹Collisions of cosmic rays with molecules in the upper atmosphere give rise to an associated “atmospheric” neutrino flux by the same mechanism.

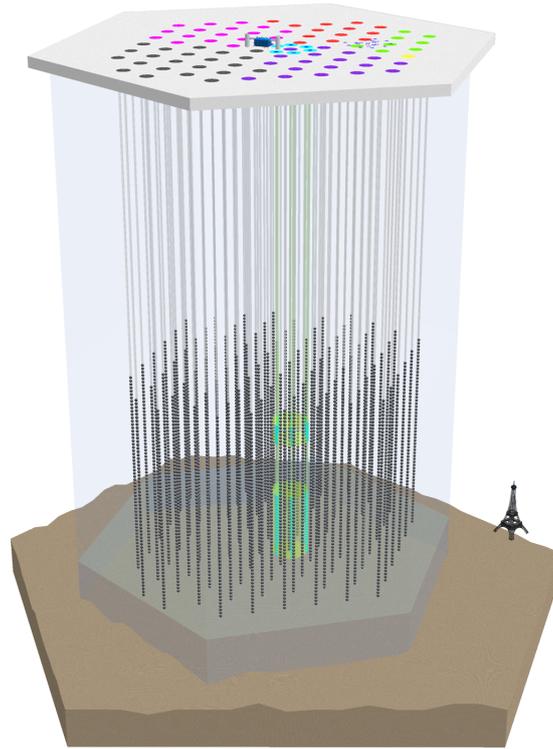


Figure 1: An illustration of the ICECUBE array, with the 324 m tall Eiffel Tower for scale comparison. The hexagonal array covers a square kilometer of surface area, and the instrumented volume of the ice is one kilometer thick. The top of the array is ~ 1500 m below the surface. The colors of the surface stations indicate the season in which each string was or will be deployed.

free of background from cosmic rays and is not limited to up-going events, making observation of the full sky possible. This work is a small part of that larger effort.

2 Detecting neutrinos with IceCube

Since neutrinos interact only weakly, they are difficult to detect. Being uncharged, they do not emit light or exert forces² electronic detectors. Accelerator experiments study the collisions of beams of particles with known momentum, producing both directly detectable particles and neutrinos. The charged particles, strongly-interacting particles, and photons will be registered in the detectors surrounding the interaction point, but the production of neutrinos must be inferred from missing transverse momentum. Observatory experiments like ICECUBE [14] take a different approach, searching for the products of weak interactions of high-energy ($E > 1$ TeV) neutrinos with matter.

2.1 Deep inelastic neutrino-nucleon scattering

Deep inelastic scattering of neutrinos on nuclei in the target material can give rise to measurable signals. All neutrino flavors participate equally in the charged-current (CC) and neutral-current (NC) reactions, given by

$$\begin{aligned} \nu_l + N &\rightarrow l + X & (\text{CC}) \\ \nu_l + N &\rightarrow \nu_l + X & (\text{NC}) \end{aligned} ,$$

where ν_l represents an incoming neutrino or antineutrino of a particular flavor (electron, muon, or tau), N the nucleon, l an outgoing charged antilepton or lepton of the appropriate flavor, and X the system of emerging hadrons. Figure 2 shows Feynman diagrams for these processes.

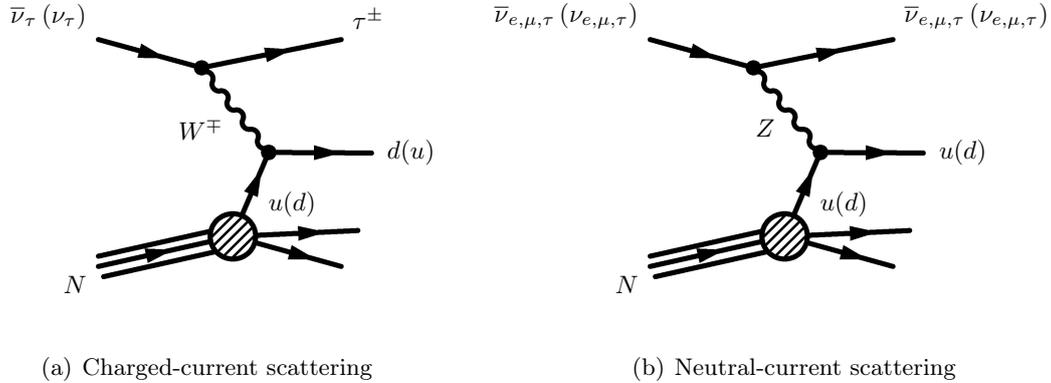


Figure 2: Feynman diagrams for deep inelastic neutrino-nucleon scattering.

Due to the universality of the weak interaction, the cross-sections for these reactions only depend on the kinematics of the reaction and the distribution of charged quarks within the nucleon. The cross-sections for an isoscalar target can be given in terms of the Bjorken scaling variables $x = Q^2/2M\nu$ and $y = \nu/E_\nu$ as [21]

$$\frac{d^2\sigma}{dx dy} = \frac{G_F^2 M E_\nu}{\pi} \begin{cases} 2 \left(\frac{M_W^2}{Q^2 + M_W^2} \right)^2 [xq(x, Q^2) + x\bar{q}(x, Q^2)(1-y)^2] & (\text{CC}) \\ \frac{1}{2} \left(\frac{M_Z^2}{Q^2 + M_Z^2} \right)^2 [xq^0(x, Q^2) + x\bar{q}^0(x, Q^2)(1-y)^2] & (\text{NC}) \end{cases} , \quad (1)$$

²With the exception of an extremely weak gravitational force that is neglected in the Standard Model [39].

where $-Q^2$ is the 4-momentum transfer between the incident neutrino and outgoing lepton, $\nu = E_\nu - E_{\{l, \nu_l\}}$ is the energy loss in the target's rest frame, M is the mass of the nucleon, M_W and M_Z are the masses of the bosons that mediate the weak interaction, and $G_F = 1.16632 \times 10^{-5} \text{ GeV}^{-2}$ is the Fermi constant. The cross-sections are linear in energy up to a few TeV where the slope decreases because of the $1/Q^2$ term in the propagator. The quark distribution functions q , \bar{q} , q^0 and \bar{q}^0 depend on $-Q^2$ and the parton momentum fraction x ; they must be extrapolated to low x by means of perturbative quantum chromodynamic calculations [12]. Total cross-sections can be obtained by integrating over all x and y .

2.2 Neutrino-electron scattering

For most energies of interest to ICECUBE, the charged-current scattering of electron antineutrinos on electrons in the target matter can be neglected. However, at a neutrino energy of 6.3 PeV, the center-of-mass energy (80.5 GeV) is large enough to produce a real W boson [20]

$$\begin{aligned} \bar{\nu}_e + e &\rightarrow W^- \rightarrow \bar{\nu}_l + l \\ \bar{\nu}_e + e &\rightarrow W^- \rightarrow X \quad , \end{aligned}$$

as shown in Figure 3, which then decays into a leptonic system $\bar{\nu}_l + l$ or hadronic system X . At this energy, the total cross-section for the reaction is much larger than the cross-sections for neutrino-nucleon scattering.

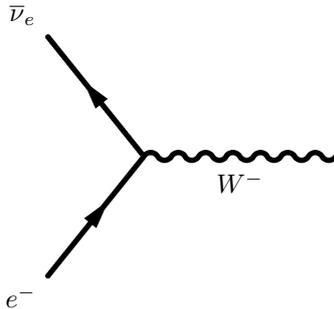


Figure 3: Feynman diagram for resonant W^- production at a neutrino energy of 6.3 PeV.

2.3 Neutrino event signatures

The signature of the above reactions is the emergence of a high-energy charged lepton or hadronic particle shower in the ice. These events can be divided into two broad classes:

- **Track-like events** arise from CC interactions of muon neutrinos. A highly relativistic muon can travel large distances through the detector before it decays, depositing energy along its path and leaving a track of light through the detector.
- **Cascade-like events** arise from the remaining interactions. A high-energy electron emerging from the interaction vertex will cause an electromagnetic particle shower, depositing most of its energy close to the interaction point. The emerging hadronic system will generate a shower as well.

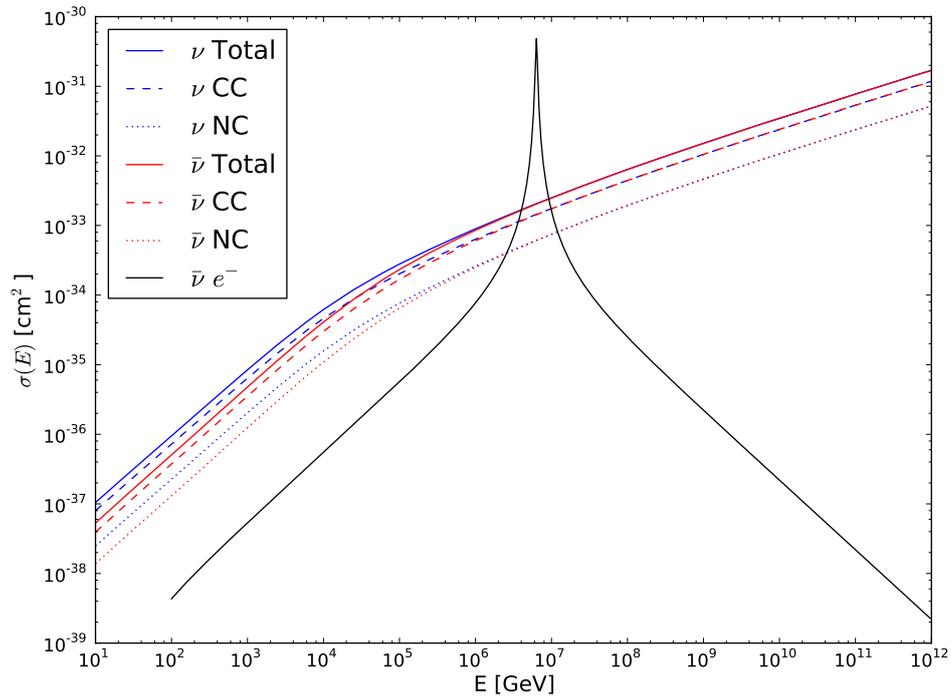


Figure 4: Neutrino-nucleon and antineutrino-electron scattering cross-sections as a function of neutrino energy from [33] based on data from [21]. From bottom to top at low energy, the cross-sections are for $\bar{\nu}$ NC, ν NC, $\bar{\nu}$ CC, $\bar{\nu}$ total, ν CC, and ν total. The resonance peaked at 6.3 PeV is the antineutrino-electron resonance.

While tracks are relatively easy to reconstruct, cascades require more elaborate methods. This work focuses on problems specific to the cascade channel.

2.4 The Cherenkov Effect

Charged particles moving through an optical medium with a velocity $v = \beta c$ greater than the local speed of light c/n emit Cherenkov radiation concentrated at the Cherenkov angle $\theta_C = \cos^{-1}(1/n\beta)$ as illustrated in Figure 5. The energy lost to radiation per unit track length can be obtained by integrating the Frank-Tamm formula over all frequencies f [24]

$$-\left(\frac{dE}{dx}\right)_{\text{rad}} = \frac{2\pi\alpha}{c_{\text{vac}}} \int_{\beta n(f) \geq 1} \left(1 - \frac{1}{\beta^2 n^2(f)}\right) df \quad , \quad (2)$$

so that the amount of Cherenkov light produced by all particles in the ice is primarily a function of the track length of the charged particles in the event.

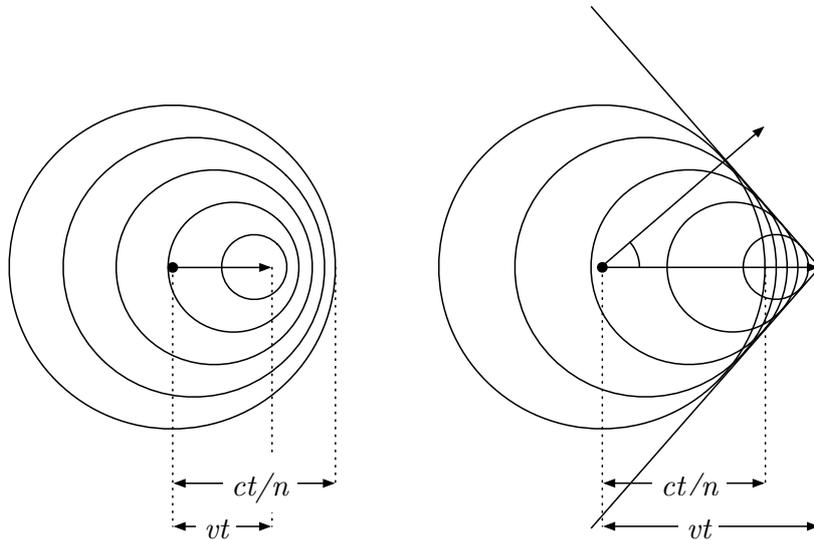


Figure 5: Illustration of the Cherenkov effect by Huygens' construction. When a charged particle travels at $\beta = 0.5$, the spherical wave fronts are only slightly concentrated along the direction of travel. At $\beta = 1$ with $n = 1.33$, however, the wave fronts interfere constructively to produce a cone of light at an angle of 41° to the direction of the charged particle.

2.5 Physics of showers

The light signature of a neutrino-induced cascade depends on the distribution of charged particle tracks within the shower. The following sections will briefly discuss the physics of cascade development and the empirical models that can be used to describe a shower's shape.

2.5.1 Electromagnetic showers

As a high-energy electron passes through matter, it is deflected by the electric fields of electrons and nuclei in the target and radiates photons through bremsstrahlung. The bremsstrahlung photons can then produce an electron-positron pair in the electric field of an atom; the electron and positron will similarly radiate bremsstrahlung photons and produce more pairs. In a simplistic model, one can assume that the electron loses half of its energy to the bremsstrahlung photon at every stage. The average distance a particle travels before its energy is reduced by $1/e$ is called the radiation length X_0 . The shower continues until the energies of the electrons in the shower drop below the critical energy E_C where energy loss due to ionization dominates over bremsstrahlung. The total track length of the shower is then

$$X = X_0 \ln \frac{E}{E_C} . \quad (3)$$

In ice, the typical values for radiation length and critical energy are $X_0 \sim 40$ cm and $E_C \sim 80$ MeV [25]. In this simple model, the length of the cascade scales with the logarithm of the energy, while the total number of electrons and hence the total track length depend linearly on the energy. Detailed simulations have shown the same behavior for realistic models of cascade development as well [37, Figures 7.48 and 7.52]. The longitudinal energy deposition of an electromagnetic cascade can be described by a Gamma distribution

$$\frac{dE}{dx} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)} , \quad (4)$$

where $t = x/X_0$ is the distance along the shower axis in units of the radiation length. Fits to simulations of electromagnetic cascades in water [37] give values of $a = 2.03 + 0.604 \ln(E_0/\text{GeV})$ and $b = 0.633$ for the shape parameters of the longitudinal distribution. The maximum of a 100 TeV electromagnetic cascade occurs 5 m from the neutrino interaction vertex.

2.5.2 Hadronic showers

The hadronic shower resulting from the struck nucleon is more complicated. The shower consists mostly of protons, neutrons, and charged and neutral pions. While the nucleons and charged pions live long enough to interact further and produce more hadrons, the neutral pions decay almost immediately to pairs of photons that can initiate electromagnetic sub-cascades. This electromagnetic leakage continues until the energy of the hadronic interactions falls below the threshold for π^0 production [19]; this happens sooner for lower-energy hadronic showers.

The total Cherenkov light yield from a hadronic shower is the sum of contributions from the charged hadrons and π^0 -induced electromagnetic cascades. The hadronic component includes uncharged neutrons that contribute no light at all, and the number of neutral pions produced depends on the shower energy. This makes the light yield from a hadronic shower both smaller than that from an electromagnetic shower and no longer simply proportional to the energy of the hadron that initiated the cascade.

The longitudinal profile of a hadronic cascade depends on both the nuclear interaction length and the electromagnetic interaction length in the detector medium. In ICECUBE simulation and analyses to date, the shower profile has been assumed to be indistinguishable from that of an

electromagnetic cascade to within the positional resolution of the detector; this assumption is discussed in more detail in [25].

2.6 Light propagation in Antarctic ice

While the IceCube detector has a large active volume, it is sparsely instrumented. The strings are spaced on a hexagonal grid 125 m apart, with the DOMs on each string 17 m apart. Even though the glacial ice at the South Pole is very pure, it contains small concentrations of dust and bubbles that scatter and absorb photons before they can be registered at the detector. Figure 6 shows measurements of the scattering and absorption coefficients of the ice at different depths. These coefficients are the inverses of the scattering and absorption lengths, respectively. The dust deposits are concentrated at certain depths, giving the ice a distinct layered structure [15]. The Cherenkov signature from relativistic charged particles in the ice is smeared out in time and space by scattering, and absorption limits the effective range from which useful information can be gathered about an event.

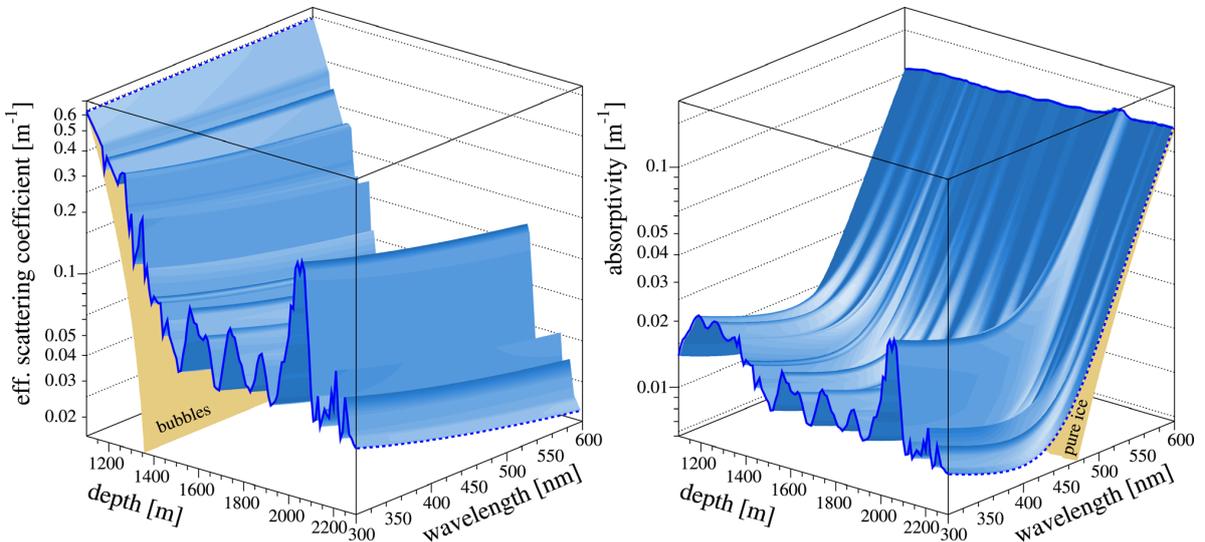


Figure 6: Wavelength and depth dependence of the scattering and absorption coefficients in South Pole ice from [15]. The IceCube coordinate $z = 0$ is located at a depth of 1946 m.

2.7 Light detection and readout

The IceCube detector relies on photomultiplier tubes (PMTs) to convert Cherenkov photons emitted by relativistic charged particles in the ice to electric signals. Each PMT is housed in a Digital Optical Module (DOM), which digitizes and time-stamps the PMT signal before transmitting it to the data acquisition hub on the surface. The operating principles of the PMT and the functions of the DOM will be explained in the following sections.

2.7.1 Photomultiplier tubes

A schematic representation of a PMT is shown in Figure 7. A PMT consists of a photocathode and several dynodes housed in an evacuated blown-glass tube. The *photocathode* is a thin

metallic coating on the inside surface of the face of the tube. When a photon strikes the photocathode, it may eject an electron with energy $hc/\lambda - \phi$, where h is the Planck constant, λ is the wavelength of the photon, and ϕ is the work function, a characteristic of the photocathode. This is the photoelectric effect. The probability that an incident photon will eject an electron is the *quantum efficiency* and depends on the wavelength and angle of incidence of the photon [30]. The Hamamatsu R7081-02 PMT used in IceCube has a quantum efficiency of 25% for normally-incident 390 nm photons [35]. An external voltage divider creates an electric potential difference between the photocathode and the first dynode, causing the ejected electron to accelerate along the associated electric field towards the dynode. When it strikes the dynode, it knocks ~ 10 – 20 electrons out of the metal of the dynode [35]; these electrons also accelerate towards the next dynode stage along the electric field created by the voltage divider and knock out further electrons. This process is repeated at every dynode stage, converting a single photoelectron into a pulse of $\sim 10^7$ electrons and giving a measurable current.

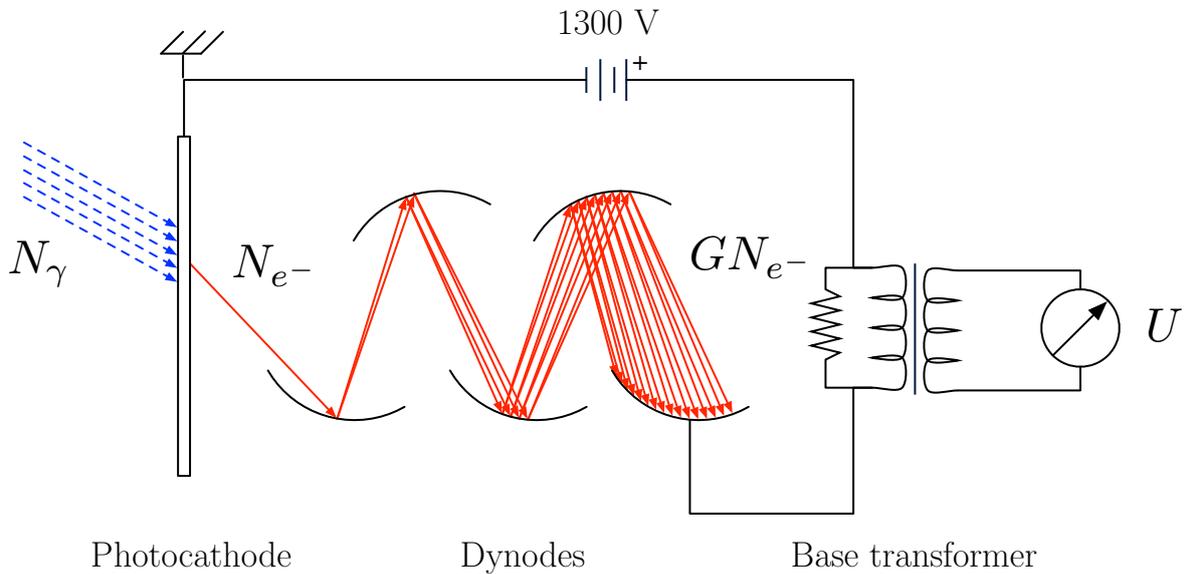


Figure 7: A schematic representation a photomultiplier tube in the configuration used in the IceCube DOMs. The voltage divider circuitry for the dynodes has been omitted for purposes of clarity. N_γ photons strike the photocathode, ejecting N_{e^-} electrons. These are accelerated in the electric field of the dynodes, each knocking out more electrons each time it strikes a dynode stage. On average, N_{e^-} photoelectrons create a pulse of GN_{e^-} electrons, where G is the nominal gain of the PMT. The last dynode stage is inductively coupled to the readout amplifier.

2.7.2 The Digital Optical Module

The basic unit of the IceCube detector is the Digital Optical Module (DOM), a pressure-resistant glass sphere containing a 10-inch PMT and readout electronics. Figure 8(a) shows a drawing of a DOM. The PMT is enclosed in a μ -metal cage to shield it from the earth’s magnetic field, which would otherwise distort the paths of electrons traveling from the photocathode to the first dynode. The cage and PMT face are bonded to the pressure sphere with room-temperature vulcanizing gel, which provides mechanical stability and enhances light transmission between the glass of the sphere and the PMT. The readout electronics are located on a circuit board around

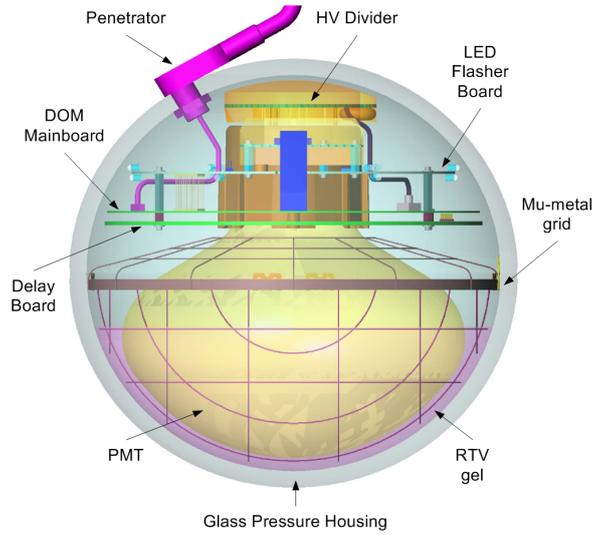
the neck of the PMT; power and communications lines run to the surface and to neighboring DOMs on the string through the penetrator at the top of the sphere.

Figure 8(b) shows a block diagram of the DOM motherboard. The output of the PMT base transformer is split into three signal paths. The first, the trigger line, is routed through a discriminator that provides a trigger signal to the DOM firmware running on the field-programmable gate array (FPGA). To avoid generating data for noise signals (thermal ionization in a dynode, decays of radioactive isotopes in the housing glass, etc.), the DOM only triggers on local coincidence. If the PMT output voltage rises above the discriminator threshold, the DOM firmware sends coincidence signals to the DOMs above and below it on the string. If the neighboring DOMs also register an above-threshold signal, indicating a light source in the ice rather than DOM noise, the DOM reads out the PMT signal over the second signal path.

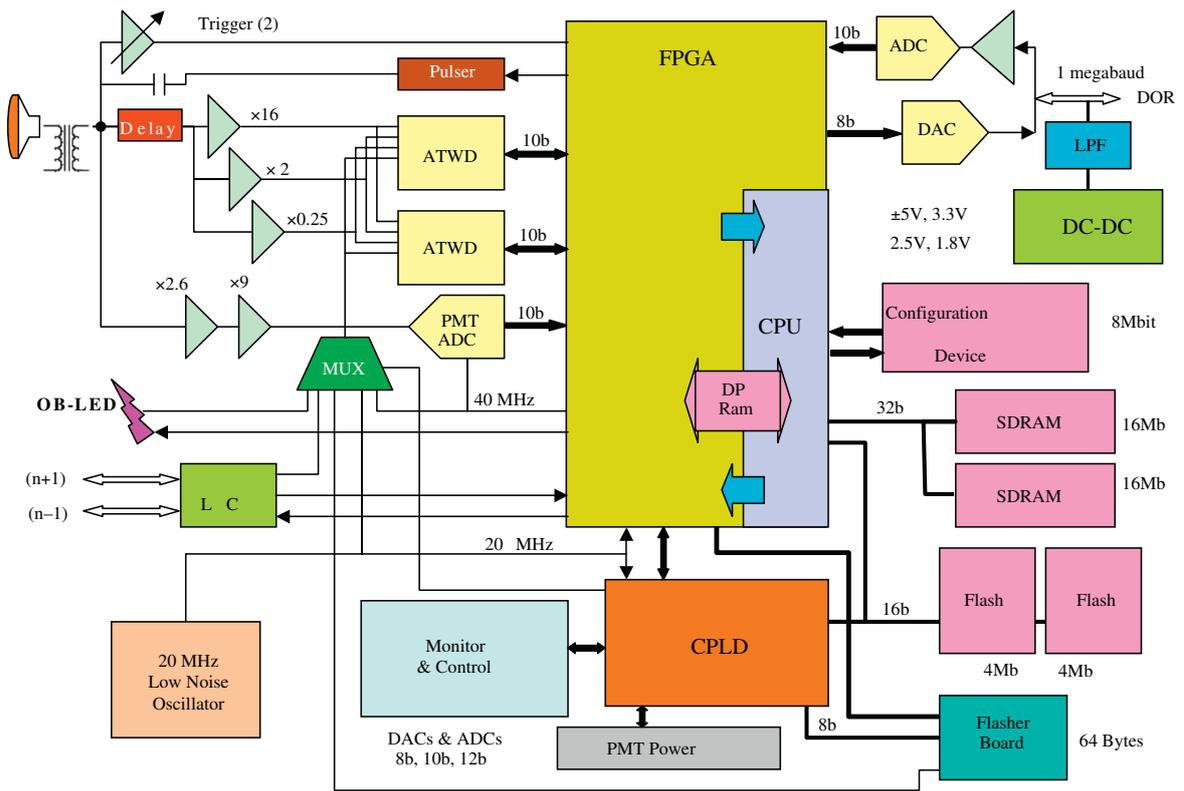
The second signal path is routed through a 75 ns delay line to allow time for a trigger decision. The delay line is read out by three amplifiers of different gains for maximum dynamic range. Each of the amplifiers is connected to a channel of custom-designed chip called the Analog Transient Waveform Digitizer (ATWD). Once trigger discriminator crosses threshold, the amplifier output voltage is sampled for ~ 3.3 ns on each of a bank of 128 capacitors. If local coincidence leads to a trigger decision, the voltage stored on each of the capacitors is digitized using a set of 128 10-bit ramping analog-to-digital converters (ADCs). The digitization process takes 29 μ s; a second ATWD is available to capture signals during this dead time.

The third signal path is connected through a pulse-shaping amplifier to a flash ADC (fADC) that continuously samples the PMT waveform every 25 ns. This ADC provides information about signals that last longer than the ATWD time window, up to 6.4 μ s, albeit with much coarser timing resolution.

The series of ADC counts (128 for each of the three ATWD channels and 256 for the fADC) and a time-stamp are transmitted digitally to the DOM hub controller on the surface. The IceCube data acquisition (DAQ) software packages this information into a data structure called a `I3DOMLaunch`. An event consists of series of `I3DOMLaunches`, one for each DOM that triggered, along with information about the overall trigger decision. The first step of reconstruction is to convert these data structures into information about the instantaneous PMT cathode current. A more detailed discussion of the ICECUBE data acquisition system can be found in [16].



(a) The IceCube Digital Optical Module (DOM).



(b) A block diagram of the DOM motherboard.

Figure 8: The IceCube DOM and DOM motherboard from [16].

3 The IceCube cascade simulation chain

Reconstructions and analysis techniques must be tested on data where the true event parameters are known. Since it is impossible to create neutrino-induced cascades under controlled conditions as with a test beam³, the next best option is to simulate the data.

The task of a detector simulation is to reproduce the physics discussed in Section 2 from the time an incident neutrino enters the Earth’s atmosphere to the data readout. The simulation should reflect all physics processes that have a significant influence on the data, converting event parameters into readout data. The task of reconstruction is the inverse, producing event parameters from readout data. If reconstructions are to be tested with simulated data, an understanding of the simulation and its limitations is necessary in order to understand the performance of a reconstruction algorithm. The following is a description of the ICECUBE ν_e simulation as currently used in large-scale production.

3.1 The IceTray software framework

The simulation and data processing software used in ICECUBE is based on a custom software framework called ICETRAY [3]. In this framework, the data belonging to an event are packaged in a Physics Frame. Each *frame* contains one or more named *frame objects*. A processing script consists of *services*, or program units that provide commonly used functions (such as a service to query a database for PMT calibration constants), and a chain of *modules*, which are program units that directly generate or manipulate data. Each frame is passed down the module chain in turn; a module can add or remove objects from the frame. In this way, tasks are separated into manageable units, and simulation and real data processing/reconstruction can be implemented in the same software framework. Each stage of the simulation discussed here is implemented as an ICETRAY module. The sequence of modules used to create the simulated data are shown in Figure 9.

3.2 Event generation

The NEUTRINOGENERATOR module [2] generates an isotropic flux of neutrinos of a given flavor⁴ according to a given power-law energy distribution. For a given neutrino, the module simulates its propagation through the Earth until it reaches the detector’s active volume. Once inside the active volume, the neutrino is forced to interact; the interaction probability (see Section 2.1) is used to calculate a statistical weight for the event. This weight can later be used along with the power-law index to convert the quantities related to the bulk properties of a simulated data set (such as number of events passing a filter criterion) to those expected for real data. NEUTRINOGENERATOR adds a structure called an I3MCTree to the frame which contains the parameters of all the energy-loss events along the neutrino’s path up to the final interaction. This is used as the input for the next stage of the simulation.

The I3MCTree contains the physics events that generate signals in the detector, for example a neutral-current scattering on a nucleon. The development of the resulting hadronic shower

³Even though there is no test beam, there are calibration light sources in the ice. Each DOM is equipped with LED “flashers” for timing calibration that can serve as vaguely cascade-like light sources in the sense that they are localized and transient [16].

⁴Here, only electron-neutrinos are considered. Other generator programs are used to simulate the background of atmospheric muons for use in realistic filter studies.

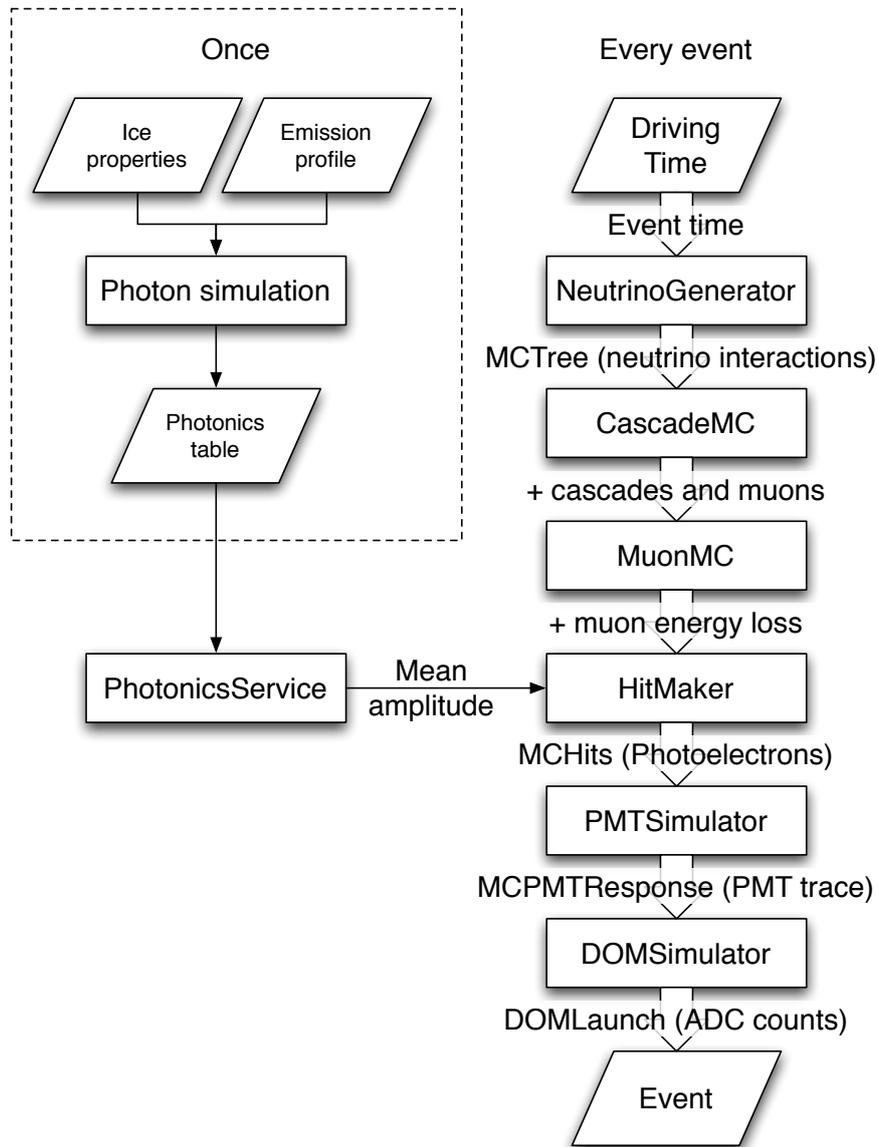


Figure 9: A schematic representation of the data flow in the ICECUBE detector simulation. The light propagation simulation is only done once; the results are tabulated for later use. The event simulation, on the other hand, is repeated every time. For each event, NEUTRINOGENERATOR generates a neutrino and propagates it through the earth to the detector, where it is forced to interact. CASCADEMC simulates the development of electromagnetic and hadronic cascades, adding light sources for each energy-loss event. MUONMC simulates energy loss by any muons produced in hadronic cascades. HITMAKER queries PHOTONICSSERVICE for the mean amplitudes tabulated for each energy-loss light source to generate photoelectrons at the photocathode of each simulated PMT. PMTSIMULATOR simulates the amplification and time delay of the photoelectrons, yielding a PMT trace. Finally, DOMSIMULATOR simulates the pulse-shaping and digitization electronics on the DOM motherboard to produce a structure of ADC counts for each DOM.

is simulated by the cascade development module, `CASCADEMC` (CMC) [2]. It does this by spreading point-like Cherenkov light emitters along the direction of the hadronic shower every few radiation lengths according to the longitudinal profile discussed in Section 2.5.1. Muons resulting from hadronic decays are also added to the particle tree; their propagation (in particular, stochastic energy losses and decay) is then simulated by the corresponding `MUONMC` [2]. Electromagnetic cascades are simpler except at high energies, where the size of stochastic losses and relativistic suppression of pair-production cross-sections make it necessary to simulate cascade development numerically (see [33] for a more detailed discussion).

At this point, the frame contains a tree representing the parent-child relationships of all light-producing energy-loss events in the detector. The next step is to simulate the propagation of photons from the sources to the DOMs.

3.3 Light propagation: Photonics

Photons traveling from light sources in the ice to the DOMs are delayed and diffused by scattering in the ice. Knowledge of the nature of these distortions is critical for both simulation and reconstruction. Since the ice at the South Pole is not homogenous, there is no simple analytic expression that can describe the time-delay of photons from an anisotropic source throughout the detector volume. This is particularly critical for cascades, because any directional information would be contained in a pronounced emission at the Cherenkov angle.

It is thus necessary to simulate the propagation of photons through the ice numerically. `PHOTONICS` [27] is a software package written for this purpose that simulates the propagation of photons through an inhomogenous medium like the Antarctic ice. In the interest of computational efficiency, photons are not tracked directly for each simulated physics event; rather, photons are propagated for various light sources and the results of the simulation are tabulated. Light distributions for `ICECUBE` event simulation are then drawn from these tables. This has the added advantage that the same tables can be used to construct a reconstruction algorithm that uses realistic information about scattering and absorption in the ice.

Figure 10 shows the geometry of the `PHOTONICS` tables. The geometry is cylindrical with the light source at the center; the fourth dimension is time. The inputs to the simulation are an angular emission profile such as Figure 7.44 from [37], the absorption and scattering coefficients of the ice layers, the depth and polar angle of the source, and the angle-dependent effective photocathode area of the DOMs. Each photon is injected into the tracking volume at ($l = 0, \rho = 0$) with an angle and wavelength drawn from the emission profile. It is then tracked in straight-line segments between scattering events; the length of these segments is determined by the local scattering coefficient. At each scattering, the photon is assigned a statistical weight representing the probability that the photon would survive to that point without being absorbed; this probability is updated according to the local ice properties. Each time the tracked photon crosses a cell of the table geometry, its weight is added to the bin at the appropriate delay time t_d , which is given by

$$t_d \equiv t - t_{\text{direct}} = t - \frac{n_{\text{ref}}}{c} \sqrt{l^2 + \rho^2} \quad , \quad (5)$$

the difference between the photon arrival time and the direct propagation time.

The bins of the table store the differential photon flux through the bin volume, binned at each delay time and normalized to the total weights of all the tracked photons; when the DOM

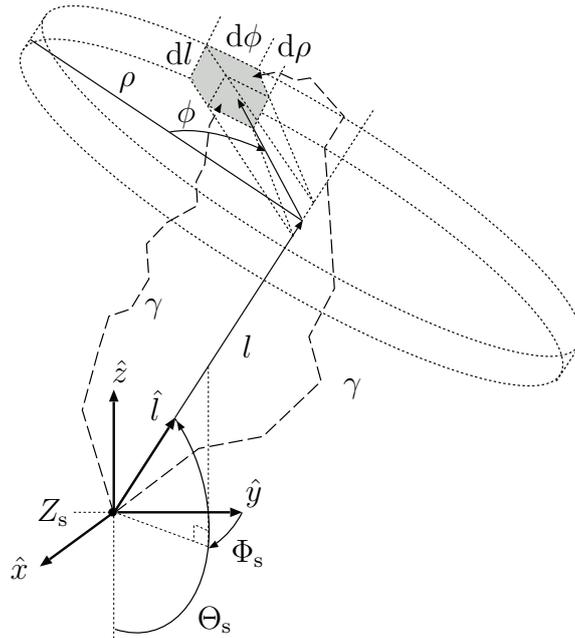


Figure 10: The PHOTONICS coordinate system, taken from [27]. Each table is binned in cylindrical coordinates with the source at the center ($l = 0$, $\rho = 0$). As each photon is tracked through the volume, its statistical weight, representing the probability that the photon would reach that point in the ice without being absorbed, is added to the bin. Each table represents a particular orientation of the source with respect to the ice layers, that is, a depth z_s in the ice and polar angle θ_s . Due to the assumed planar symmetry of the ice layers, a single table can be used for all azimuthal angles ϕ_s and (x,y) positions of the source, reducing the dimensionality of the problem significantly.

efficiencies are taken into account, this gives the differential photoelectron yield at delay time t_d for a DOM at the point (l, ρ) from the light source. The expected number of photoelectrons for a cascade event of a particular energy can be obtained via

$$n = \mu \Lambda_\gamma L(E) \quad , \quad (6)$$

where μ is the value stored in the table bin, Λ_γ is the mean number of Cherenkov photons produced in the sensitivity range of the DOMs per unit track length (32440 m^{-1}), and $L(E)$ is the integrated track length of the shower as a function of energy (4.37 m/GeV for electromagnetic cascades [37]).

The PHOTONICS tables are used in ICETRAY through PHOTONICSSERVICE [1]. This interface loads the tables into memory and provides the mean photoelectron yield at a given delay time and DOM position from a light source somewhere in the ice, handling the details of converting between coordinate systems and calculating the total photon yield for a source type with a given energy. Mean amplitudes for DOM positions corresponding to points between the centers of table bins are calculated by linear interpolation between neighboring bins in each dimension. Given the mean photoelectron yield at the photocathode, the simulation can generate photoelectrons in the PMTs.

3.4 Hit generation

The HITMAKER module [2] simulates photoelectron emission for each energy loss event in the frame by using PHOTONICSSERVICE to sample from the photon arrival time distribution for each DOM. Each electron emission event is characterized by its arrival time, weight (number of emitted electrons), and source⁵; this information is encapsulated in an `I3MCHit`. The `I3MCHitSeries` produced by the module is effectively a convolution of the PHOTONICS arrival-time distribution with the transit-time spread and pre/late/afterpulse distribution of the PMT. This serves as the input to the PMT simulation.

3.5 PMT simulation

The PMT simulation implemented in the PMTSIMULATOR module [2] accounts for the amplification characteristics of the PMT. Each `I3MCHit` is converted to an amplified charge according to PMT charge response distribution, spread out in time according to the PMT pulse shape, and converted to a voltage using the nominal gain of the PMT and effective impedance of the coupling transformer. A saturation curve is applied to the output waveform to account for the measured non-linearity in response to large instantaneous photocathode currents. The `I3MCPMTResponse`

⁵ The majority of the PMT signal consists of “normal” pulses from photoelectrons propagating directly from the photocathode through the dynode stages. Large pulses are accompanied by a small number nuisance events from other sources: pre-pulses, late pulses, and afterpulses. Pre-pulses result from a photon striking the first dynode stage directly, releasing a small number of electrons and resulting in a small pulse $\sim 30 \text{ ns}$ before the main pulse. Late pulses $\sim 60 \text{ ns}$ after the main pulse are consistent with electrons backscattering from the first dynode stage before striking again and being amplified in the usual fashion. Very late pulses (300 ns to $11 \mu\text{s}$ after the main pulse) are consistent with ionization of residual gases in the PMT by electrons accelerating between the dynode stages. The positively charged ions are accelerated back towards the photocathode, where they eject more electrons that lead to an afterpulse. HITMAKER generates these auxiliary hits using parameterizations derived from laboratory measurements of the PMT pulse time distribution. Details of these measurements are given in [35].

produced by this module represents the voltage at the input amplifier connected to the PMT coupling transformer⁶.

3.6 DOM simulation

The DOMSIMULATOR module [2] implements the delays and distortions introduced by the input amplifiers between the PMT coupling transformer and the ATWD and fADC as well as noise in the digitization electronics. For each I3MCPMTResponse where the voltage crosses the trigger threshold, DOMSIMULATOR produces a I3DOMLaunch containing a timestamp and a series of ADC counts for each channel of the ATWD as well as the fADC. The series of I3DOMLaunches represents the same kind of data that can be obtained from the real detector, and is the input to the first stage of reconstruction.

⁶The droop in the transformer output voltage in response to extended pulses varies between different types of DOMs, so it is implemented in DOMSIMULATOR.

4 Cascade reconstruction

The most basic element of a reconstruction is the event hypothesis, that is, a model of the physical event underlying the data. The hypothesis determines which parameters should be used to describe the event; a specific reconstruction is just a way of determining that parameters of the model that best describe the data. This section will describe the parameters of the cascade model as well as some elements of low-level processing that are common to all event reconstructions.

4.1 Event parameters

For cascade-like events, one naturally chooses a cascade hypothesis. A cascade-like event can be described by the location of the cascade vertex (x , y , and z), the time of the event (t), the orientation of the cascade (polar angles θ and ϕ), and the visible energy of the cascade (E). The meaning of these parameters will be explained here.

4.1.1 Coordinate system

The location and orientation of a cascade event are defined with respect to the ICECUBE coordinate system, a right-handed coordinate system with its origin close to the geometrical center of the detector at a depth of 1946 m below the surface. The y axis points “north” along the Prime Meridian, while the z axis points vertically upward towards the surface. The direction of a particle in the ice is defined by the vector pointing back towards its source; particles traveling directly downwards relative to the surface have a zenith angle of 0 while up-going particles have a zenith angle of π . The azimuthal angle is defined with respect to the x axis as usual. Figure 11 shows a sketch of the 5 ICECUBE spatial coordinates. Positions are measured in meters and angles in radians⁷.

4.1.2 Timing

Trigger times are defined in nanoseconds since the beginning of the calendar year, universal time coordinate (UTC)⁸. The time of the (reconstructed) underlying physics event is defined relative to the trigger time. Details of clock synchronization with the Global Positioning System (GPS) are given in [16].

4.1.3 Visible energy

In the model of cascade development given in Section 2.5, the light output of a cascade is a function of the total length of the charged-particle tracks in the event over which Cherenkov radiation is emitted. Since no distinction is made between electromagnetic and hadronic cascades, the visible energy is a proxy for the total Cherenkov track length. The visible energy of a cascade is defined as the energy of an electromagnetic cascade with the same light output; the

⁷This is different from the PHOTONICS coordinate system, which uses degrees internally.

⁸The maximum granularity of the master clock is 0.1 ns.

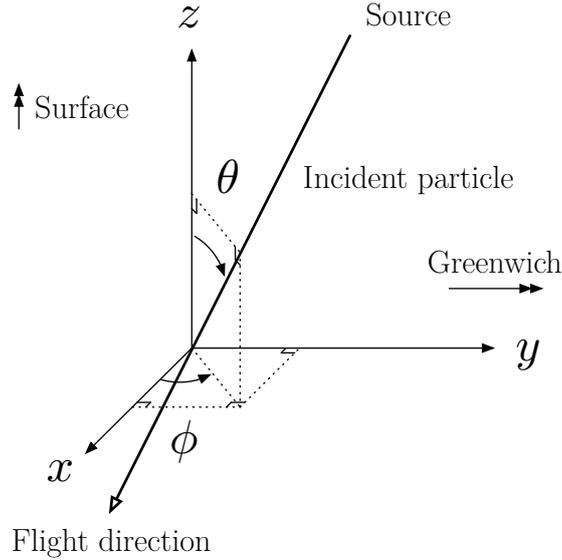


Figure 11: The ICECUBE coordinate system. Since the direction vector points back towards the source of the particle, the zenith angle is inverted relative to zenith angle in conventional spherical coordinates.

total energy released in a hadronic cascade will be larger than the visible energy [25]. Energies are measured in gigaelectron volts (GeV)⁹.

4.2 Low-level reconstruction

The first stage of reconstruction is to convert the raw data, consisting of structures of ATWD and fADC counts, into physical quantities that can be compared to an event hypothesis. This proceeds in two stages: DOM calibration and pulse reconstruction.

4.2.1 DOM calibration

The raw event data consist of counts from each ATWD gain channel as well as the fADC. These must be converted to voltages for physics analysis while correcting for known instantaneous effects of the electronics. These corrections are implemented in the DOMCALIBRATOR module [1]. For each ATWD bin, the module takes the equivalent input voltage for the highest-gain channel and subtracts the pedestal voltage if needed, falling back to lower-gain channels if the higher-gain channel is saturated. This ensures that the digitized waveform has the highest possible voltage resolution while staying within the dynamic range of the ADC. The fADC counts are also converted to equivalent voltages. The waveform bins are shifted earlier in time to account for the gain-voltage-dependent PMT transit time. Furthermore, the module attempts to correct the temperature-dependent droop of the PMT output transformer for extended pulses [34]. The output of the module is a series of calibrated waveforms that reflect the idealized output of the PMT transformer. Information about the photon arrival time distribution can be extracted from this waveform.

⁹1 GeV=10⁹ eV= 1.60217646 × 10⁻¹⁰ joules.

4.2.2 Pulse reconstruction

The calibrated waveform itself does not reflect the photon arrival time distribution, since the shaping amplifiers in front of the digitization electronics introduce distortions to the measured waveform. In particular, the pulse shapes for the ATWD and fADC are different. For this reason, it is useful to unfold the true PMT waveform from the digitized data. The PULSEEXTRACTOR module [5] implements a Bayesian unfolding algorithm for this purpose.

Since the pulse-shaping characteristics of the amplifiers have been measured, it is in principle possible to represent the pulse shaping as a matrix operation that transforms the time series of PMT voltages \vec{v} into a time series of ATWD inputs \vec{a} :

$$a_i = \sum_j A_{ij} v_j. \quad (7)$$

The ATWD input is a discrete convolution of the PMT waveform with the amplifier pulse shape. In this framework it is generally not possible to extract the PMT time series

$$v_i = \sum_j A_{ij}^{-1} a_j, \quad (8)$$

because the matrix A is not generally invertible¹⁰. An approach based on Bayes' Theorem (see Section 5.2) developed in [8] avoids this difficulty by instead extracting the most probable distribution \vec{v} from the time series \vec{a} given knowledge of the pulse-shaping function $P(a_i|v_j)$. The inverse conditional distribution is given by

$$P(v_i|a_j) = \frac{P(a_j|v_i)P_0(v_i)}{\sum_k P(a_j|v_k)P_0(v_k)}, \quad (9)$$

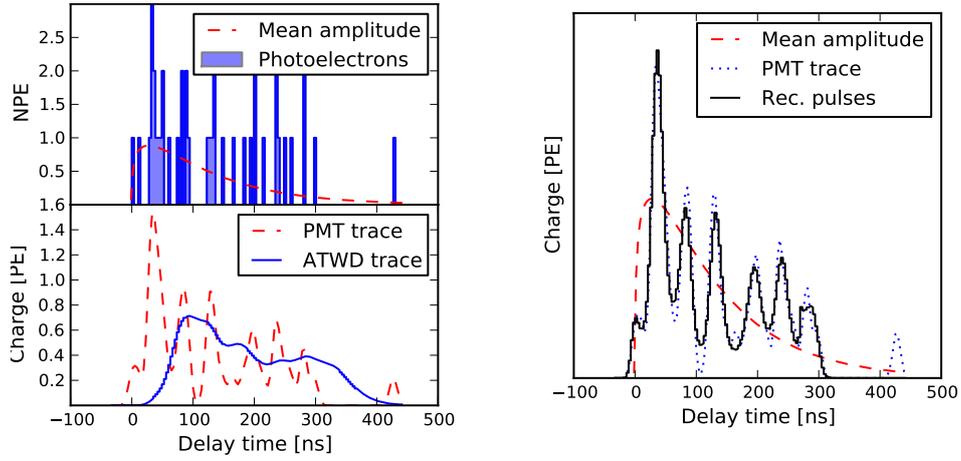
where $P_0(v_i)$ is the prior distribution or best guess at the true distribution of \vec{v} ¹¹. This can be used to estimate the PMT voltage in bin i ,

$$\tilde{v}_i = \sum_j a_j P(v_i|a_j). \quad (10)$$

This algorithm can be iterated to obtain better and better approximations to \vec{v} . The prior distribution for the next iteration is then just the normalized $\tilde{\vec{v}}$:

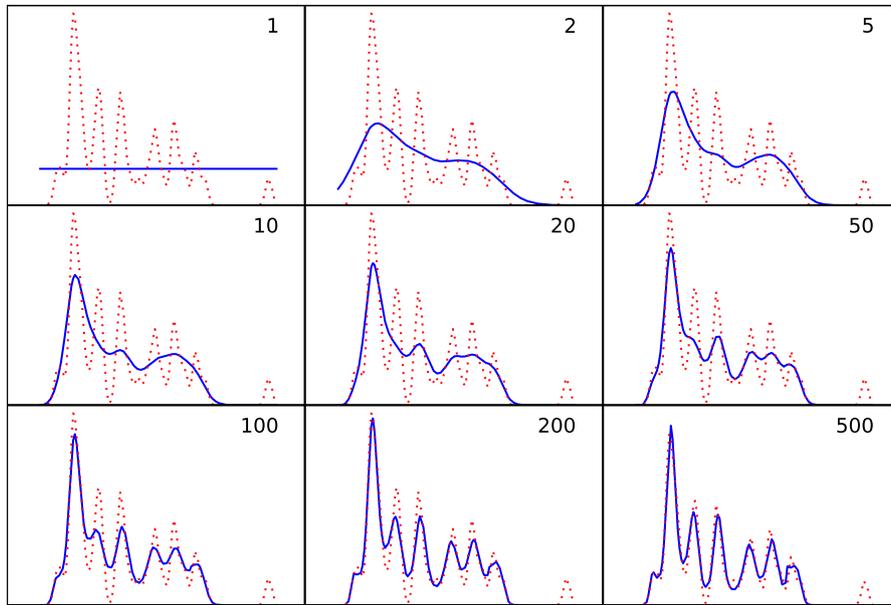
$$P_0(v_i) = \frac{\tilde{v}_i}{\sum_j \tilde{v}_j} \quad (11)$$

Figure 12 illustrates the use of this algorithm to reconstruct PMT pulses from the ATWD trace¹². (a) shows how the ATWD trace relates to the PMT trace, photoelectron arrival times, and true photon arrival time distribution (mean amplitude). (c) shows how the reconstructed



(a) Inputs to pulse extraction

(b) Result of the unfolding



(c) Iterations of the unfolding algorithm

Figure 12: An illustration of how PMT pulses are reconstructed from the ATWD trace. The data are from a toy simulation. In (a), 42 photoelectrons are drawn from the delay-time distribution (given by the mean amplitude) for an isotropic source 60 m from a DOM. The delay time distribution is calculated using the empirical model developed in [32]. The PMT trace is obtained by convoluting the delay time distribution with the PMT pulse shape, described by a normal distribution; the ATWD trace is obtained by convoluting the PMT trace with the ATWD pulse shape, described by a log-normal distribution. The same ATWD pulse shape is used as the kernel for the Bayesian unfolding, shown in (c). The time shift is removed after only a few iterations, while many more are needed to recover the exact waveform. The reconstructed pulses shown in (b) are a good approximation to the PMT trace and can be compared to the mean amplitude predicted for an event hypothesis as part of a reconstruction.

pulse distribution converges to the PMT trace under repeated applications of Equation (10). (b) shows the final reconstructed pulses along with the mean amplitude and ATWD trace.

Since the sampling windows of the ATWD and fADC overlap, some finesse is required in combining waveforms from the two devices without double-counting. Here, only the ATWD is used.

4.3 First-guess algorithms and filtering

The calibrated waveform and reconstructed pulses are the basis for many *first-guess* algorithms, fast reconstructions that use the geometry and timing of the event to make rough guesses at the event parameters. The average position of the hit DOMs weighted by their collected charge, for example, can be used to estimate the position of the cascade vertex, and the total charge can be compared to a parameterization derived from simulations to estimate the energy. Since such comparisons can be made relatively quickly, they are also used to separate neutrino-induced cascades that warrant more resource-intensive reconstruction from background events.

Good first-guess algorithms and efficient filtering are a prerequisite for a real cascade analysis, since they can reduce the required computing effort to a manageable level. The results of the first-guess algorithm are used as the input or *seed* to more intensive reconstructions. This work builds on [29] and focuses solely on the problems of more advanced reconstructions, skipping the problems of seeding and filtering. Further details on these topics are given in [25, Sections 5.1, 7.2, 7.3] and [29, Section 4.3]. Given a good seed, one can proceed to the most advanced reconstruction technique, based on the method of maximum likelihood.

¹⁰Hence the term “smearing.”

¹¹In the case of complete ignorance, the best choice is a uniform prior $1/N$, where N is the length of the time series.

¹²The convolution kernel used here is a log-normal distribution with parameters that approximate the response function of the ATWD amplifier. The same kernel was used for the unfolding the data, all of which were produced with a toy simulation.

5 Likelihood methods

The reconstruction algorithm studied in this work is based on the method of maximum likelihood. This section will introduce the theoretical framework for this method and formulation of the likelihood before introducing methods of estimating the inherent uncertainty of such a reconstruction.

5.1 Statistical inference

The problem of event reconstruction is one of statistical inference. Given a measurement from the detector and a model of the detector response to different events, one must extract information about the underlying event. The simplest case to imagine is one in which the detector response is unambiguous, that is, it is impossible that two different classes of event produce the same detector response. In this case, the event may be classified with absolute certainty. Difficulties arise when the detector response is ambiguous, that is, when two different events produce similar signatures¹³. In these more realistic cases, one must move from statements of absolute certainty to probabilistic statements. Instead of implied certainty, each result will be tempered by a “degree of belief.” What is required, then, is a formal framework for describing this degree of belief.

5.2 Bayes’ Theorem

Bayes’ Theorem provides a method for computing conditional probabilities. To introduce this theorem, it is first necessary to introduce the concept of probability. Given a set S with subsets A, B, \dots , the probability P is defined by the following axioms:

1. $P(A) \geq 0 \forall A \in S$
2. $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$
3. $P(S) = 1$.

The conditional probability “A, given B” is defined [17, Equation (31.3)] as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (12)$$

Rewriting this and using the fact that the intersection $A \cap B$ is the same as $B \cap A$, one obtains Bayes’ Theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \quad (13)$$

By requiring that all probabilities be properly normalized, one can rewrite this as

¹³An extreme example of this is an event in ICECUBE where light is only detected at a single string. If one considers only the hits, the light distribution has complete azimuthal symmetry and events on opposite sides of the string are indistinguishable.

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_i P(B|A_i)P(A_i)}. \quad (14)$$

To apply this to event reconstruction, one defines a function $P(\vec{x}|\vec{\lambda})$ that describes the probability of the detector response \vec{x} to an event parameterized by $\vec{\lambda}$. One can then apply Bayes' Theorem to define a probability for the event hypothesis $\vec{\lambda}$:

$$P(\vec{\lambda}|\vec{x}) = \frac{P(\vec{x}|\vec{\lambda})P(\vec{\lambda})}{\sum_i P(\vec{x}|\vec{\lambda}_i)P(\vec{\lambda}_i)} \quad (15)$$

The term $\sum_i P(\vec{x}|\vec{\lambda}_i)P(\vec{\lambda}_i)$ enforces the normalization of $P(\vec{\lambda}|\vec{x})$.

5.3 Maximum likelihood estimators

The function $P(\vec{\lambda})$ gives the independent probability that an event with parameters $\vec{\lambda}$ would occur in the experiment. It is also called the ‘‘prior distribution,’’ since it describes the experimenter’s knowledge of the process under study before the measurement is made. In the case of complete ignorance of the distribution of the parameters $\vec{\lambda}$, one can use a uniform prior distribution to construct a ‘‘Likelihood Function’’ for the event data given by \vec{x}_m

$$L(\vec{\lambda}) = cP(\vec{x} = \vec{x}_m|\vec{\lambda}) \quad (16)$$

where c is a constant that absorbs any factors that depend only on the data [7, Section 7.1]. It is important to note that the likelihood function cannot be interpreted as an absolute probability, since it is not necessarily normalized. The usefulness of the likelihood function comes from the Likelihood Axiom, as formulated by Edwards:

Within the framework of a statistical model, *all* the information which the data provide concerning the relative merits of hypotheses is contained in the likelihood ratio of those hypotheses on the data, and the likelihood ratio is to be interpreted as the degree to which the data support the one hypothesis against the other [10, Section 3.3].

For each event, the likelihood function can be treated as a function of the event hypothesis $\vec{\lambda}$ only and be maximized analytically or numerically with respect to the hypothesis parameters to yield the hypothesis of maximum likelihood. This the set of parameters $\vec{\lambda}'$ best supported by the data \vec{x}_m and the detector description given by $P(\vec{x}|\vec{\lambda})$.

5.4 A likelihood function for cascades

This work is based on the formulation of a likelihood function for cascade-like events developed and tested in [29]. It relies on the assumption that since the propagation of each Cherenkov photon through the ice is independent, the charge registered by a DOM in any given time window follows a Poisson distribution

$$P(n|\mu) = \frac{\mu^n}{n!} e^{-\mu} \quad , \quad (17)$$

where n is the number of detected photoelectrons and μ is the mean number of photoelectrons (mean amplitude) predicted by PHOTONICS in that time window for a point-like electromagnetic cascade. The probability of obtaining a given event signature given a parameterization of μ can then be written as

$$L = \prod_{\substack{\text{hit DOMs} \\ o}} \prod_{\substack{\text{bins} \\ i}} \frac{\mu_{o,i}^{n_{o,i}}}{n_{o,i}!} \exp(-\mu_{o,i}) \prod_{\substack{\text{unhit DOMs} \\ o}} \exp(-\mu_o) \quad , \quad (18)$$

where the time bins of DOMs with no charge are grouped together to speed up the calculation. The mean amplitude is a function of the DOM position $(x, y, z)_{\text{DOM}}$, readout time window t_r , and event hypothesis parameters $(x, y, z, t, \theta, \phi, E)$ (see Section 4.1). If L is a good description of the detector response, then the parameters of the event hypothesis that maximize L are the best approximation to the true event parameters.

5.5 Error estimates from the likelihood function

In general, the parameters of maximum likelihood will not coincide with the true event parameters. Just as a given event can result in a number of different detector responses, each with a certain probability, so a given detector response can arise from a number of different events, each with a certain probability¹⁴. Thus, one expects the location of the likelihood maximum in the parameter space to fluctuate around the true parameters. What is needed, then, is a way to estimate the resolution of the reconstruction, or the expected difference between the parameters of maximum likelihood and the true event parameters.

One of the simpler methods of calculating the uncertainty of a reconstruction is to compare the results to simulation. To do this, one simulates a large sample of events using parameters from some physically motivated spectrum. After the reconstruction has been applied to the data generated by this simulation, one can calculate the ‘‘residual’’ for each event for each parameter λ_i :

$$\delta_{\lambda_i} = \lambda_{i,\text{reconstructed}} - \lambda_{i,\text{simulated}} \quad (19)$$

The error or resolution of the reconstruction for each parameter λ_i can then be estimated from the variance of δ_{λ_i} over all events in the sample. While this method can be used to characterize the average resolution for the events in the simulation sample, it cannot yield any information about the inherent resolution of the detector for an individual event.

This information can be extracted from the likelihood function itself. Since the likelihood function represents a statistical model of the detector response, its shape around the maximum holds information about the inherent error of the estimate. Given two hypotheses $\vec{\lambda}_i$ and $\vec{\lambda}_j$, the likelihood ratio is defined as

¹⁴There is a subtle difference between the two statements that is also encapsulated in the difference between equations (15) and (16). While $P(\vec{x}|\vec{\lambda}')$ describes possible fluctuations of the detector response to an event $\vec{\lambda}'$, $L(\vec{\lambda})$ describes the degree to which different hypotheses $\vec{\lambda}$ are statistically compatible with the data under the assumption that all values of $\vec{\lambda}$ are equally likely *a priori*. While the measured data are a result of the combined effects of fluctuations in detector response and the true spectrum of event parameters, the fluctuations described by the likelihood function with a flat prior are not constrained by the true spectrum.

$$\Delta L(\vec{\lambda}_i, \vec{\lambda}_j) = \frac{L(\vec{\lambda}_i)}{L(\vec{\lambda}_j)} = \frac{P(\vec{x} = \vec{x}_m | \vec{\lambda}_i)}{P(\vec{x} = \vec{x}_m | \vec{\lambda}_j)} . \quad (20)$$

It is particularly useful to fix the denominator at the maximum likelihood $L(\vec{\lambda}')$. Edwards uses this ratio to define the *support* or relative log-likelihood function:

$$S(\vec{\lambda}) \equiv \ln \left(\frac{L(\vec{\lambda})}{L(\vec{\lambda}')} \right) = \ln L(\vec{\lambda}) - \ln L(\vec{\lambda}') . \quad (21)$$

The support for the hypothesis of maximum likelihood is 0, while the support for all other hypotheses is negative. $-S(\vec{\lambda})$ is a measure of how much better the data support the hypothesis $\vec{\lambda}'$ than $\vec{\lambda}$ [10].

The error of the maximum likelihood estimate can be defined as the volume of parameter space¹⁵ V_{λ} around the maximum $\vec{\lambda}'$ where the support is greater than some fixed number¹⁶. Under the assumption of equal *a priori* parameter probabilities, the $\vec{\lambda}$ within this volume are also statistically compatible with the maximum likelihood estimate. Thus, the true event parameters should also lie within this volume. An estimate of the size of this volume is thus also an estimate of the typical residual for that event.

5.6 Modelling the detector response

The principle of maximum likelihood places no restrictions on the form of the likelihood function. To a first approximation, any function that has a maximum near the true event parameters provides a useful estimate of the true event parameters. The suitability of a likelihood function in this sense can be evaluated by the bulk residual method mentioned above. The Poisson-likelihood for cascades certainly satisfies this requirement, as shown in [29].

For an estimate of the error from the likelihood function itself, however, more is required. An arbitrary function with a maximum near the true event parameters cannot be expected to reflect the variance in detector responses that gives rise to the reconstruction error. For the likelihood-based error estimate to be useful, the likelihood function must be a good statistical model of the detector response.

One basic requirement for a “good statistical model” is that the conclusions drawn from comparisons of hypotheses by their likelihoods in the model not violate common sense. In particular, the support for the set of true event parameters should not be so low as to reject it completely in favor of another hypothesis. If the support at the set of true parameters is well below the cutoff used to calculate the error volume, then it lies outside the error volume and the size of the error volume is no longer a good indicator of the expected residual.

This can be checked using the detector response as modelled in the ICECUBE simulation. With the simulation, one has both the true event parameters and the detector response in hand and can observe how well the likelihood function models the fluctuations in the detector response.

¹⁵In the gaussian limit with uncorrelated parameters, this is the usual n -dimensional ellipse with principal axes $\sigma_1, \sigma_2, \dots, \sigma_n$.

¹⁶Again in the gaussian limit, the conventional 2σ error volume corresponds to 2 units of support or a log-likelihood ratio of 2.

Very low support for the true event parameters is an indication that the model contains approximations that significantly distort the likelihood function¹⁷, and that correcting the model can lead to better parameter estimates (i.e. support for the true parameters) and accurate estimates of the reconstruction residual.

5.7 Calculating the error volume

There are a number of possible methods for calculating the error volume, most of which amount to an estimate of the covariance matrix of the likelihood function, defined as

$$V_{ij} = \langle (\lambda_i - \langle \lambda_i \rangle)(\lambda_j - \langle \lambda_j \rangle) \rangle \quad (22)$$

If the individual components of the measurement are statistically independent, then the likelihood function converges to a gaussian in the large-sample limit¹⁸. In this limit, the series expansion of the logarithm around the maximum,

$$\ln L(\vec{\lambda}) = \ln L(\vec{\lambda}') + \frac{1}{2} \sum_{i,j} \frac{\partial^2 \ln L(\vec{\lambda}')}{\partial \lambda_i \partial \lambda_j} (\lambda_i - \lambda'_i)(\lambda_j - \lambda'_j) + \mathcal{O}(|\vec{\lambda} - \vec{\lambda}'|^3) \quad , \quad (23)$$

can be terminated after the second-order term [7, Equation 7.6.3]. Exponentiating this, one obtains an expression of the form

$$L(\vec{\lambda}) = k \exp \left[-\frac{1}{2} (\vec{\lambda} - \vec{\lambda}')^T V^{-1} (\vec{\lambda} - \vec{\lambda}') \right] \quad , \quad (24)$$

which is a multivariate normal distribution with covariance matrix V . The inverse of the covariance matrix can then be approximated by

$$(V^{-1})_{ij} = -\frac{\partial^2 \ln L(\vec{\lambda}')}{\partial \lambda_i \partial \lambda_j} \quad (25)$$

The elements of the covariance matrix can be calculated either by evaluating the second-order derivatives of $\ln L$ or by fitting an n -dimensional paraboloid centered at $\vec{\lambda}'$ to $\ln L$. Both of these approaches require knowledge of either the analytic form of the likelihood function or the typical scale of variations in the function, which is what the method is intended to calculate in the first place. Without knowledge of the appropriate scale, it is necessary to scan the parameter space; this can become problematic in high dimensions.

The most general method is to evaluate the variance directly using the N -dimensional integral

$$V_{ij} \equiv \int (\lambda_i - \bar{\lambda}_i)(\lambda_j - \bar{\lambda}_j) L(\vec{\lambda}) d^N \vec{\lambda} \quad , \quad (26)$$

¹⁷For example, a model that does not account for PMT saturation leads to a likelihood function that over-penalizes saturated waveforms, since it expects Poisson fluctuations inversely proportional to the mean amplitude. A detailed discussion of these approximations follows in Section 7.

¹⁸This is a consequence of the central limit theorem, see [7, Equation 5.11.3]

where the mean along each dimension is given by

$$\bar{\lambda}_i \equiv \int \lambda_i L(\vec{\lambda}) d^N \vec{\lambda} \quad . \quad (27)$$

Since the integral runs over all dimensions, this method will always be sensitive to dependencies between the parameters. This can be an advantage for the reconstruction of local events like cascades, where the position of the vertex can effect the most likely angle of photon emission.

When the analytic form of the likelihood function is not known, this integral must be evaluated by numerical methods. Conventional methods of polynomial approximation can become problematic in large dimensions. In Monte-Carlo integration, one generates a sample of points whose distribution is proportional to the likelihood function. The variance of this sample will converge to the true variance of the likelihood function as the number of samples approaches infinity. The same sample can be used to approximately evaluate any integral over the likelihood function. This study focuses on use of a Markov chain to efficiently generate such a sample and methods of using the resulting distribution to find and correct flaws in a likelihood-based reconstruction.

6 Markov Chains

This work uses a Markov-Chain Monte Carlo technique to calculate the error volume of a likelihood-based reconstruction. The following sections will introduce the basic theory of Markov chains as well as the implementation used in this work.

6.1 Theory

A Markov chain¹⁹ is a sequence of random variates $\{X^{(t)}\} = X^{(0)}, X^{(1)}, X^{(2)}, \dots$ that has the Markov property, that is, that the influence of $X^{(0)}, X^{(1)}, X^{(2)}, \dots$ on the distribution of $X^{(n+1)}$ is mediated entirely by the value of $X^{(n)}$. The steps of the chain t can be interpreted as times. A Markov chain can be specified by stating the initial distribution of X^0 (the distribution of starting points of the chain) and the conditional distributions of $X^{(n+1)}$ given $X^{(n)}$ (the two-point transition probabilities $T_n(x, x')$). The transition probabilities can be used to calculate the distribution of variates at time $t = n + 1$ given the distribution of variates at time n :

$$p_{n+1}(x) = \sum_{\tilde{x}} p_n(\tilde{x}) T_n(\tilde{x}, x) \quad . \quad (28)$$

This procedure can be used recursively to find the distribution of the chain at all later times. The *invariant* or *equilibrium distribution* of a Markov chain is one that persists forever once reached, that is, becomes independent of time. The distribution π is invariant if for all n ,

$$\pi(x) = \sum_{\tilde{x}} \pi(\tilde{x}) T_n(\tilde{x}, x) \quad . \quad (29)$$

It is often useful to construct a Markov chain that generates samples from a given distribution. This requires that the given distribution be the invariant distribution of the chain. The distribution and transition probabilities should satisfy the condition of *detailed balance*, meaning that any transition from x to x' should be reversible, that is, have the same probability of the transition from x' to x :

$$\pi(x) T(x, x') = \pi(x') T(x', x) \quad . \quad (30)$$

In addition, the chain should be *ergodic*. This means that the chain should converge to its equilibrium distribution as $n \rightarrow \infty$, regardless of the starting point. It can be shown that a fairly general class of Markov chains is ergodic. If the transition probabilities and equilibrium distribution satisfy the condition

$$\nu = \min_x \left[\min_{x': \pi(x') > 0} \left(\frac{T(x, x')}{\pi(x')} \right) \right] > 0 \quad , \quad (31)$$

then the chain converges to π regardless of the starting point, and the rate of convergence is bounded by

$$|\pi(x) - p_n(x)| \leq (1 - \nu)^n \quad . \quad (32)$$

¹⁹This section follows the definitions of [31, Section 3.3].

A proof is given in [31]. In effect, the convergence of the chain is bounded by the most inaccessible region of the parameter space. Accordingly, the bound obtained in this manner is often too weak to be useful, especially on continuous, infinite spaces. It is thus sometimes necessary to resort to heuristics to gauge the convergence of the chain.

The end product of the Markov chain should be a sequence of nearly independent samples from π . The amount of computational time required to produce a set number of independent samples depends on three factors.

- **Transition probability calculation** The transition probability $T(x, x')$ must be calculated in every iteration. This limits the number of samples that can be produced in a given time.
- **Autocorrelation length** Since the sample produced in each step is obviously dependent on the previous step, it takes some time to produce another sample that is approximately independent. The intermediate samples must be discarded. The autocorrelation length is a number of steps between uncorrelated samples.
- **Burn-in length** Similarly, the samples should not depend on the starting point of the chain. The burn-in length is the number of the samples from the start of the chain that must be discarded before the samples are independent of the starting point. This can be different from the autocorrelation length at other points in the chain.

6.2 Implementation

The preceding discussion was general to time-reversible, ergodic Markov chains. The remaining discussion will focus on the Metropolis Algorithm, a specific implementation of a Markov sampler.

6.2.1 The Metropolis Algorithm

The Metropolis Algorithm, detailed in [28]²⁰, was the first use of Markov chain techniques for Monte Carlo integration. The specific application was a calculation of the equation of state of an interacting gas, which involves an integral over a large phase space. Since large regions of the phase space made only small contributions to the integral, the authors proposed Algorithm 1 for generating samples from the most significant regions of the phase space. At each step of the chain, a small change to the state Δx is drawn from a *proposal distribution* $S(x, x + \Delta x)$ to generate a *proposed transition* $x_n \rightarrow x' = x_n + \Delta x$. The probability of the transition $x_n \rightarrow x'$ is taken as $\min(1, \pi(x')/\pi(x_n))$, where π is the distribution being sampled from. Transitions to states of higher probability are always taken, while transitions to states of lower probability are taken with a probability given by the odds ratio $\pi(x')/\pi(x_n)$.

The Metropolis algorithm is ergodic as long as the proposal distribution $S(x, x + \Delta x)$ is symmetric with respect to x and $x + \Delta x$ and non-zero over the parameter space [31]. The choice of proposal distribution is left open and can have a strong effect on the performance of the sampler. A normal distribution is a fairly general choice; it is symmetric and non-zero everywhere.

²⁰Properly, this name refers only to the original version in which $\pi(x')/\pi(x)$ was calculated according to the distribution of energy states in the canonical ensemble as $\exp[-(E(x') - E(x))/k_B T]$. This method was later generalized by Hastings [23] to problems outside of statistical mechanics, that is, general target distributions π . In this form, it is sometimes known as the Metropolis-Hastings algorithm.

Algorithm 1 The Metropolis Algorithm

```
for all  $n$  do  
   $x' \leftarrow x_n + \Delta x$  //  $\Delta x$  is drawn from the proposal distribution  
   $p \leftarrow \min(1, \pi(x')/\pi(x_n))$   
   $z \leftarrow \text{random}()$  //  $z \in (0, 1]$   
  if  $z \leq p$  then  
     $x_{n+1} \leftarrow x'$   
  else  
     $x_{n+1} \leftarrow x_n$   
  end if  
end for
```

The scale (variance) of the proposal distribution should be chosen to minimize the autocorrelation time of the chain. Heuristics from studies such as [22] suggest that for d -dimensional π which are approximately multivariate normal, a multivariate normal proposal distribution maximizes efficiency with a scale factor of $2.38/\sqrt{d}$, accepting roughly 1 out of every 4 proposed jumps.

6.2.2 Python implementation

The Metropolis Algorithm itself is not difficult to implement; this work uses a custom implementation written in the Python programming language, the core of which is not much longer than the pseudocode in Algorithm 1. The sampler was tested with known distributions in small numbers of dimensions to gauge its performance and develop effective tests of convergence. Figure 13 shows one of these tests, using a 3-dimensional gaussian as a target distribution. The following paragraphs will discuss the relevant features of this implementation.

Likelihood calculation Since this implementation is intended to sample from the likelihood distribution of a cascade event, the expensive calculation of the probability π is kept inside the ICETRAY framework. The interface is provided by a Python function that takes a list of parameters and returns a log-likelihood, which is then used to calculate an odds ratio for a proposed jump.

Proposal distribution The proposal distribution is a multivariate normal distribution with a diagonal covariance matrix. Each jump can optionally be taken in only a subset of the dimensions to improve the probability of accepting a jump. The proper scaling factors for the elements of the matrix were estimated by running the sampler with a large number of iterations on test data using a proposal distribution with uniform variance. The mean variance of the likelihood function for the data as a function of event charge was parameterized and used as the variance of the proposal distribution.

Autocorrelation and thinning The samples used to estimate attributes of the target distribution should be uncorrelated. Before it is used, the chain must be *thinned* to remove correlations between the samples²¹. The *autocorrelation time* is the mean number of steps that it takes for

²¹An example of large-scale correlation is easy to see in the first 500 samples of Figure 13. The samples along the “trail” from the starting point to the region of the maximum are obviously not independent.

the chain to generate another independent sample. The *autocorrelation function* of a variable x is defined as [18, Appendix B]

$$C_{xx}(t) \equiv \langle x_{(i)}x_{(i+t)} \rangle - \langle x_{(i)} \rangle \langle x_{(i+t)} \rangle \quad . \quad (33)$$

$C_{xx}(0)$ is just the variance of x . The normalized autocorrelation function²²

$$\Gamma_{xx}(t) \equiv \frac{C_{xx}(t)}{C_{xx}(0)} \quad (34)$$

typically decays as $e^{-t/\tau_{exp}}$, where τ_{exp} is the *exponential autocorrelation time*. Another measure of the autocorrelation is the *integrated autocorrelation time*,

$$\tau_{int} \equiv \frac{1}{2} + \sum_{t=1}^{\infty} \Gamma_{xx}(t) \quad . \quad (35)$$

A sequence of N samples only contains $N/2\tau_{int}$ effectively independent samples [18]; it must be thinned by this factor to remove correlations between the samples.

The autocorrelation function can be estimated from a discrete sample as

$$\tilde{C}_{xx}(t) \equiv \frac{1}{N} \sum_{i=1}^N \left\{ \left(x_{(i)} - \frac{1}{N} \sum_{j=1}^N x_{(j)} \right) \left(x_{(i+t)} - \frac{1}{N} \sum_{j=1}^N x_{(j+t)} \right) \right\} \quad . \quad (36)$$

This summation for all $t \in [0, N-1]$ can become expensive for large N . If x is transformed to $x' = x - \bar{x}$, then this reduces to

$$\tilde{C}_{x'x'}(t) \equiv \frac{1}{N} \sum_{i=1}^N x'_{(i)}x'_{(i+t)} \quad , \quad (37)$$

which is just the discrete convolution of the time series of x' with itself. The convolution is implemented using fast Fourier transforms, reducing the $\mathcal{O}(N^2)$ computation to $\mathcal{O}(N \ln N)$. This is important when using the autocorrelation function to monitor convergence of the chain.

Determining convergence The sampler uses a relatively simple criterion to determine convergence. At certain intervals, the chain is divided into bins of several times the integrated autocorrelation length of the sampler position²³. The variance of the sample position in each dimension is calculated repeatedly, each time excluding one of these bins from the calculation²⁴. If the exclusion of any one bin changes the variance of the chain by less than some set amount (e.g. 5%), then the sample distribution is roughly independent of time and the sampler is terminated. The typical size of the jackknife bins is illustrated by the dashed vertical lines in Figure 13.

²²A good overview of serial correlation is given in [6, Chapter 2].

²³The integrated autocorrelation length is generally different for each variable. In this case, the largest integrated autocorrelation length among the dimensions of the parameter space is used.

²⁴This is known as jackknife binning [18, Appendix C].

Output The Python program runs the Metropolis sampler until the convergence criterion is satisfied. Once the sampling run is finished, it records the mean, variance, and kurtosis of the samples along each dimension of the parameter space, as well as the integrated autocorrelation length in each dimension. It also records the value of the target distribution function at the starting point of the chain and the maximum value over all points in the chain.

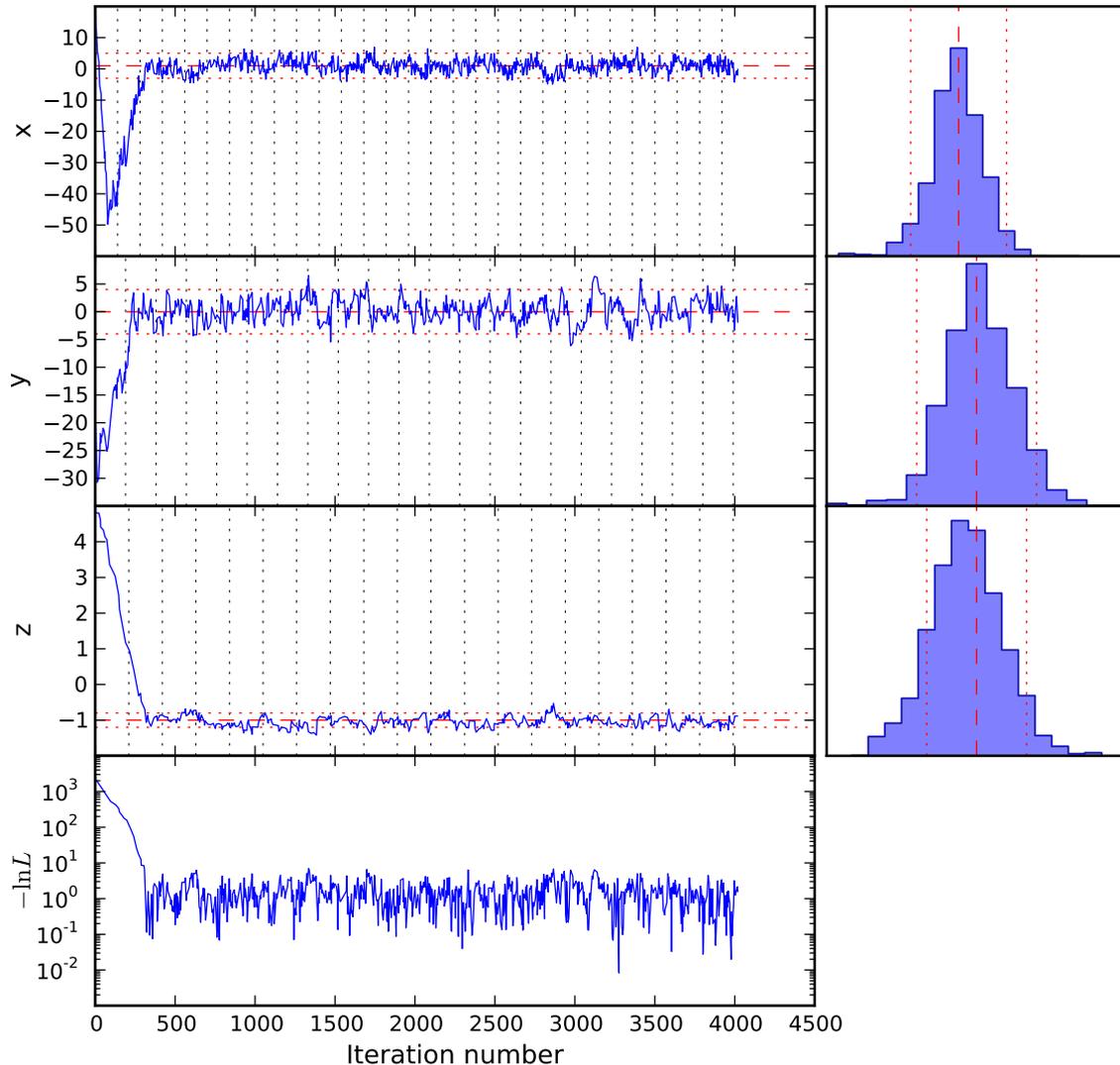


Figure 13: An illustration of a Markov sampler running in 3 dimensions. The target distribution is a multivariate normal distribution with correlation between x and z . The proposal distribution is also multivariate normal, but with independent distributions in x , y , and z and slightly larger variance. The upper three plots show the sampler position in x , y , and z in each iteration. The dashed horizontal line gives the mean of the target distribution; the dotted horizontal lines the $\pm 2\sigma$ boundaries. The histograms on the right show the marginal distribution of sampled points along each dimension. The lower plot shows the negative log-likelihood at each point; the sampler moves mostly “downhill” from the starting point (far from the maximum) until it finds a region of high likelihood, then begins to move around the maximum, probing its shape. The dashed vertical lines are spaced 10 autocorrelation lengths apart; the autocorrelation length is different in each dimension. Convergence is checked by jackknife binning with bins much wider than the longest autocorrelation length; if the variance of the samples in the chain does not change by more than 5% when any of the jackknife bins is omitted, then the chain has converged to the target distribution.

7 Application to cascade reconstruction

The Metropolis sampler can be used with the likelihood function given in Equation (18) to build a reconstruction that gives both a maximum likelihood estimate and an estimate of the error for cascade-like events in ICECUBE. This is useful, because the results from the initial study of the reconstruction’s performance on simulated data [29] showed some features that warranted direct investigation of the characteristics of the likelihood function. First, the event-by-event resolution was highly variable. Given this fact, it was only possible to quote some “core” resolution for the reconstruction that may not be meaningful for a significant subset of the events. Second, there was a correlation between the residuals in different dimensions. This suggested that the behavior of the likelihood function in a less interesting dimension, for example the depth of the cascade vertex in the ice, could significantly affect the behavior of the function in a physically interesting dimension like the angular orientation of the cascade. Both of these features indicate the need for an event-by-event measure of the resolution that takes into account the behavior of the function in all dimensions. The Markov-Chain Monte Carlo technique is ideal for this purpose.

7.1 Simplified reconstruction problem

The likelihood-based reconstruction of cascade-like events is equivalent to the numerical minimization of $-\ln L$ in 7 dimensions: the position of the cascade vertex, x, y, z ; the time of the cascade event t , the orientation of the cascade ϕ and θ ; and the cascade’s visible energy E . In general, $-\ln L$ will have more than one local minimum, and the outcome of the minimization procedure will depend on the starting point in the parameter space, called the “seed.” In real-world applications, the seeding strategy is a critical part of the reconstruction; if the seed is very far away from the true event parameters, the minimizer may become stuck in a local minimum, causing the reconstruction to fail.

These seed strategies can be extended by using an iterative fitting technique. In this method, the minimizer is started with the seed and allowed to converge to a new set of parameters that better fit the data. This set of parameters is then varied slightly and used as the seed to the next iteration of the minimizer. In this way, the minimizer can sometimes be coaxed out of local minima and towards the global minimum of the function. While useful for navigating complicated likelihood spaces, this technique is computationally expensive, and will like any minimization strategy succeed only if the minimum of $-\ln L$ is near the true event parameters.

This last question is an important one: given a perfect seed, how far will a minimizer wander from the true event parameters? The performance of the reconstruction when seeded with the true event parameters is a test of the characteristics of the likelihood function only, and can help uncover flaws in the formulation and implementation of the function that do not depend on the seeding or minimization strategy used. This study focuses only on this simplified reconstruction problem and the application of Markov-Chain Monte Carlo techniques to identify the weak points of the PHOTONICS-based reconstruction of cascade-like events in ICECUBE.

7.2 Benchmark data set and implementation

The Markov sampler was tested on a sample of ~ 16000 ν_e -induced cascade events simulated in the 80-string configuration of ICECUBE. The arrival directions of the primary ν_e were distributed

isotropically and their energies distributed according to an E^{-1} spectrum, i.e. uniformly in the logarithm of energy²⁵, with energies between 100 GeV and 1 PeV. For simplicity, the sampler was restricted to move only in 4 dimensions: the cascade vertex position x , y , z , and cascade event time, t . The cascade orientation and energy were held fixed at their true values.

For each point in the chain, the sampler records the value of the log-likelihood. Even though the Markov sampler is not technically a maximizer, it is likely to come very close to the nearest maximum of the likelihood function. The samples in the chain thus contain information about both the maximum of the likelihood function and its variance. The maxima of the likelihood function together with the starting point of the chain (the true event parameters) yield the residual distributions for each parameter.

An error can be estimated from the bulk residual distribution in several ways. The two that will be used here are the gaussian fit and the truncated variance. In the former, a normal distribution is fit to a histogram of the residual distribution. This is likely to describe the variance of the core of residuals concentrated in the peak of the distribution, but will not describe the minority of events in the tails of the distribution. In the latter method, one first estimates the sample mean and variance

$$\mu' = \frac{1}{N} \sum_{i=1}^N \delta_{z_i} \quad (38)$$

$$\sigma' = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_{z_i}^2 - \mu'^2} \quad , \quad (39)$$

then removes the residuals that differ from the mean residual by more than $3\sigma'$. The truncated variance is the variance of the remaining M samples:

$$\tilde{\sigma} = \sqrt{\frac{1}{M} \sum_{|\delta_{z_i} - \mu'| < 3\sigma'} \delta_{z_i}^2 - \left(\frac{1}{M} \sum_{|\delta_{z_i} - \mu'| < 3\sigma'} \delta_{z_i} \right)^2} \quad . \quad (40)$$

The truncation mitigates the influence of large outliers on $\tilde{\sigma}$ in a way similar to the gaussian core fit. These variances will be called *bulk errors*.

If the implementation of the likelihood function used here is a good description of the simulated measurement process, then the average variance of the likelihood function around its maximum should be similar to the bulk error calculated from the residual distribution.

First, it is useful to look at a typical residual distribution. Figure 14 shows the residual distribution in the depth (z) of the cascade vertex with a fit to a normal distribution. The error reconstruction in z is obviously not well described by a single normal distribution. In addition, there is a bias towards negative residuals, that is, the reconstructed vertex tends to lie below the true vertex.

Figure 15 shows the distribution of the estimated variance of the likelihood function in z as calculated from the variance of the sample points from the Markov chain. The mean variance of

²⁵This energy spectrum is unrealistic, but ensures an equal number of events in each decade of energy.

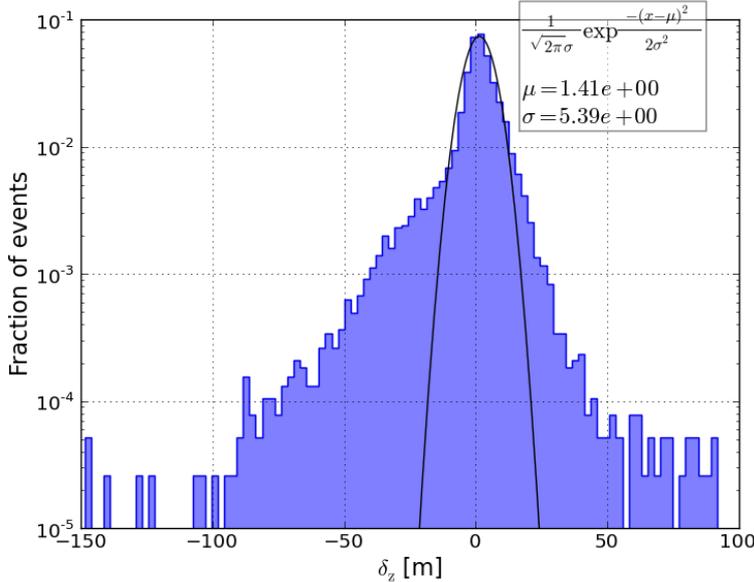


Figure 14: Bulk residual distribution of the vertex depth z for ~ 16000 simulated ν_e induced cascades weighted to an E^{-1} spectrum. The reconstruction was seeded with the parameters of the simulated cascade event. The logarithmic scale is chosen to emphasize the fraction of events whose residual is not described by the gaussian fit.

the likelihood function drops off as a function of the total charge (NPE), as one would expect for a process dominated by the statistics of photoelectron emission at the PMT photocathode²⁶.

The points in Figure 15 show the truncated variance (see Equation (40)) of the residual distribution for each range of NPE²⁷. Unlike the variance of the likelihood function, the variance of the residual distribution is almost constant, that is, dominated by systematic rather than statistical errors. In particular, the reconstruction is not significantly better for an event with 1000 photoelectrons than an event with 30. The sampler finds a very narrow likelihood peak very far from the true event parameters, indicating that the likelihood function is not a good description of the simulated measurement process for events with more than a few photoelectrons. Similar effects can be seen in the other dimensions (x , y , and t).

This behavior is reflected in the likelihood ratio between true and reconstructed event parameters²⁸, shown in Figure 16. Contrary to the expectation for a suitable likelihood model, the likelihood ratio scales strongly with the number of photoelectron emissions in the event. For events with e.g. 3000 photoelectrons, the reconstructed event parameters are on average e^{1000} more likely than the true parameters. This excessive suppression of the “correct” solution to the inference problem indicates that the likelihood function does not approximate the simulation very well, and explains why the variance of the likelihood function around its maximum does not reliably predict the residual.

These arguments show that the existing implementation of the PHOTONICS-based likelihood

²⁶Since the number of Cherenkov photons emitted is proportional to the cascade energy, an analogy can be made to sampling calorimeters, whose energy and position resolution scale as $1/\sqrt{E}$ [17].

²⁷A total recorded charge of 1000 PE usually corresponds to a cascade energy of 10–30 TeV, though fluctuations from 2 TeV to 1 PeV are possible (see Figure 34).

²⁸This is minus the support of true parameters $-S(\vec{\lambda}_{\text{true}})$, or the degree to which the maximum-likelihood hypothesis is favored over the true parameters (see Section 5.6).

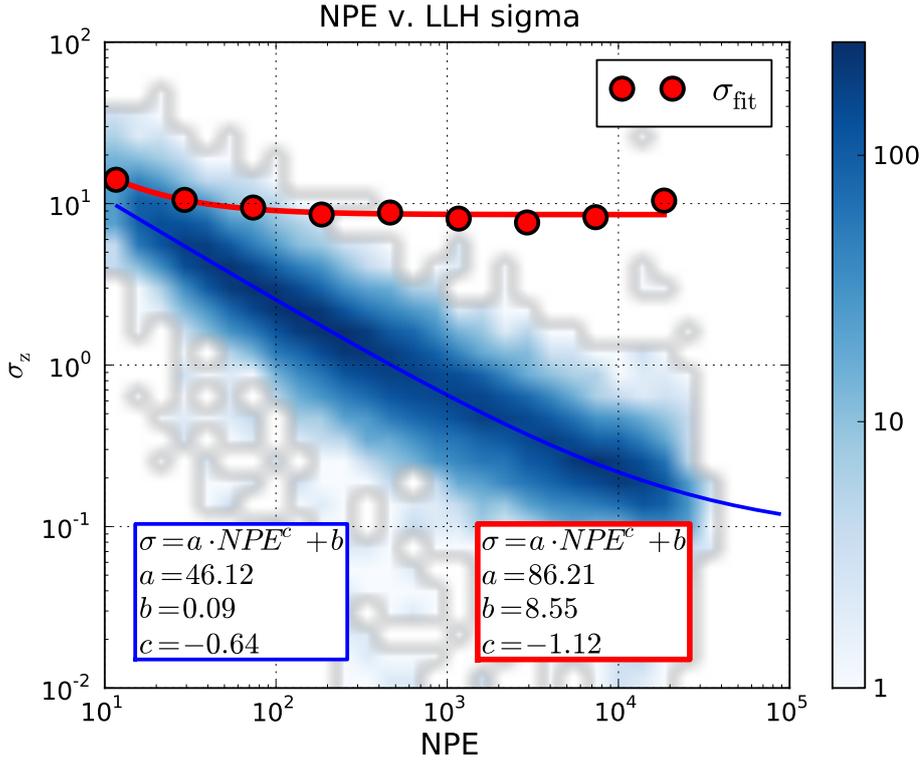


Figure 15: Resolution in vertex depth z (σ_z) as a function of total event charge (NPE) for ~ 16000 simulated ν_e induced cascades weighted to an E^{-1} spectrum. The two-dimensional histogram shows the variance of the likelihood function in z , while the points show the variance of the residual distribution for each range of NPE. In each case, the error is parameterized as a charge-dependent statistical error plus a constant systematic term. While the variance of the likelihood function falls off as a function of NPE, the slices of the bulk residual distribution are dominated by a systematic error of roughly 9 m.

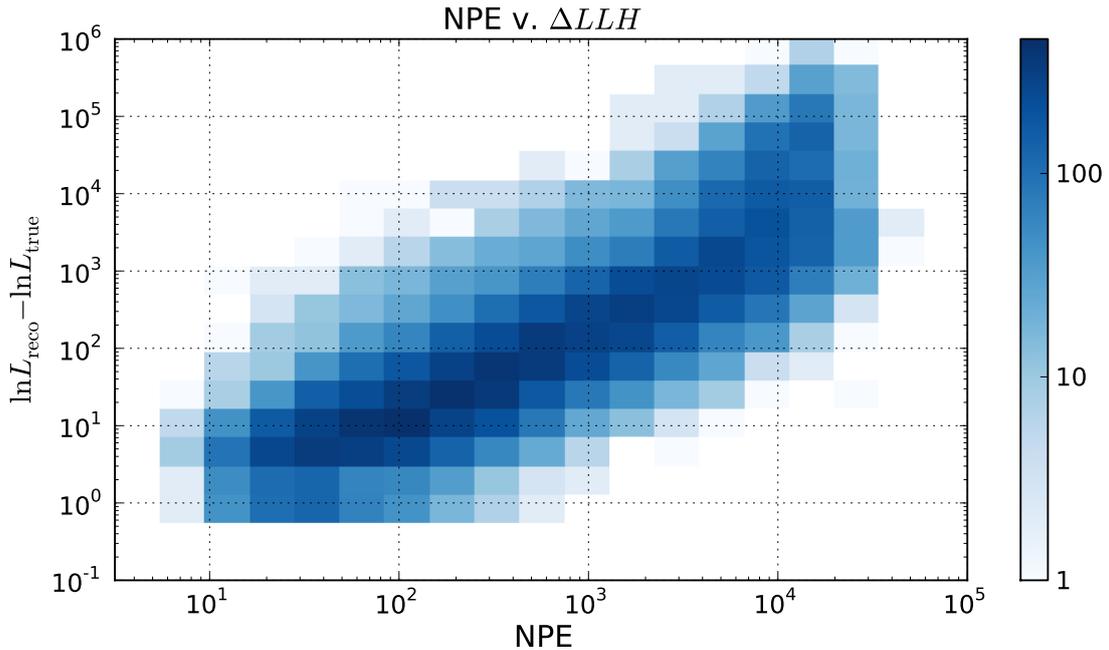


Figure 16: Logarithm of the likelihood ratio between true and reconstructed event parameters (see Section 5.6) as a function of total event charge (NPE). For bright events, the true vertex becomes extremely unlikely in the model represented by the likelihood function, indicating the likelihood does not describe these events well. By construction, the shape of the likelihood around the maximum cannot be used to estimate the residual in such cases.

function was not entirely compatible²⁹ with the standard ICECUBE cascade simulation. While there are certainly situations in which anomalies in the statistical model such as discontinuities and multiple maxima can give rise to misleading estimates as discussed in [10, Chapter 8], these concerns do not apply to a purely gaussian model. Since the likelihood function should approach this limiting case for events with many photoelectrons, the reasons for the failure of the likelihood reconstruction were not immediately apparent.

7.3 Implicit approximations in the likelihood function

The existing formulation of the likelihood function implicitly assumed that the various other fluctuations and distortions implemented in the simulation are negligible compared to Poisson fluctuations in the photocathode current. Some of these assumptions are problematic, causing the likelihood description of the detector response to fail as shown in the previous section. The most obvious of these assumptions are discussed here.

Light emission profile The PHOTONICS light propagation simulation operates using point-like light sources that can be combined to approximate extended light sources [27]. The neu-

²⁹A careful reader might object to the strength of this statement and argue that the reconstruction is in fact compatible with the simulation in the sense that it localizes the vertex depth to within 9 m. This interpretation, however, does not use the full power of likelihood methods. If all the information provided by the support function is taken in to account, one would also conclude that the vertex depth for an event with ~ 10000 photoelectrons is likely to be localized to within some tens of centimeters, which is patently false. The aim of the remaining work is to resolve such contradictions.

trino propagation simulation approximates the longitudinal evolution of an electromagnetic or hadronic shower in the ice by splitting the cascade event into point-like Cherenkov light sources distributed along the direction of the shower [33]. When a cascade event is simulated, the total light output of all the sub-cascades is determined by an energy-dependent parameterization obtained from another simulation, described in [37]. The reconstruction, however, operates on the hypothesis of a single, point-like Cherenkov light source.

Equivalence of PHOTONICS tables used in simulation and reconstruction The PHOTONICS tables contain the mean light yield at different displacements from the light source, binned in cylindrical coordinates: the parallel distance l along the source direction, the perpendicular distance ρ from the source direction, the azimuthal angle ϕ , and the delay time t . If the binning is too coarse, the tables may miss important features in the light distribution, but finely binned tables may become much too large to fit into computer memory. In simulation, one can work around this by only loading a small subset of the tables needed at the moment, since the parameters of the light-producing events are known. Reconstruction is more difficult, since the relevant table region to be queried by the minimizer may not be known beforehand. To work around this, thinned table sets are used which have fewer depths and zenith angles of the source, with linear interpolation between them. In addition, the tables used for reconstruction have a slightly different, though still linear, interpolation scheme between the bins in a single table. Interpolation artifacts could dominate the reconstruction, especially at high intensities, where large relative fluctuations around the mean are heavily penalized in the Poisson model.

PMT charge response A photoelectron ejected from the photocathode does not always produce the same integrated current at the PMT base. The nominal gain of the PMT is determined from the location of the peak of the charge response distribution for a single photoelectrons (see Figure 17). If the amplifications for multiple photoelectrons are independent, as the simulation assumes, then the multi-PE charge response can be obtained by repeatedly convoluting the single PE charge response distribution with itself. The exponential tail to low PE means that the average charge response is not the same as the nominal gain. The extension of the Poisson likelihood from natural to real non-negative numbers assumes that the fluctuations in charge response are negligible compared to the Poisson fluctuations in the photoelectron yield of the photocathode. Furthermore, it assumes that they can be accounted for entirely by scaling down the mean PE yield from PHOTONICS (μ in Equation (18)) by the mean of the SPE charge response function. This is about 0.85 for the “standard” PMT used in simulation [4].

PMT transit time and pulse shape The photoelectrons ejected from the cathode knock out additional electrons as they strike each dynode, forming a larger and larger pulse of current. The time it takes the pulse to reach the PMT base scales with the square root of the applied PMT voltage [35], and thus varies from DOM to DOM. The shape of the pulse in time is approximately a gaussian with a width on the order of 2 ns. The reconstruction assumes that this time spread can be accounted for in bulk, that is, that the PHOTONICS mean waveform is a good predictor of the PMT signal when convolved with a gaussian of the same width as the PMT pulse.

Spurious PMT pulses In particular, the reconstruction does not account for pre-pulses, after-pulses, and late pulses (see footnote on page 20). Pre-pulses in particular could be problematic, since the mean amplitude predicted by PHOTONICS is identically zero at negative delay

times. The reconstruction assumes that the mean amplitude of the recorded waveform can be predicted from the PHOTONICS mean amplitude simply by accounting for the geometric delay and mean PMT transit times.

Waveform reconstruction The ICECUBE DOM reads out the PMT base current through a transformer and series of amplifiers. This yields a captured waveform that is not the PMT signal, but the PMT signal convoluted with the amplifier pulse shape. The PMT signal can be de-convoluted from the captured waveform with an iterative unfolding algorithm such as the one implemented in PULSEEXTRACTOR [5]. The formulation of the likelihood function assumes that this process is perfect, i.e. that it accurately reconstructs the integrated current at the PMT base with no significant distortions in time or charge.

Saturation The PMT base current is proportional to the number of photoelectrons ejected from the photocathode when the latter number is small. The constant of proportionality is the nominal gain. For large photocathode currents (e.g. from a cascade very close to the DOM), the base current can be less than proportional to the number of photoelectrons. At a nominal gain of 10^7 , the PMT response is linear to within 10% for photocathode currents of up to ~ 31 PE/ns[35]. The formulation of the likelihood function assumes that the PMT nonlinearity is negligible compared to Poisson fluctuations of the number of photoelectrons ejected from the cathode. Any DOM with instantaneous photocathode current above the saturation level is likely to distort the likelihood calculation, since the PHOTONICS mean waveform does not account for saturation effects.

Zero-charge bins The Poisson likelihood works by comparing the charge in each time bin to the charge predicted by PHOTONICS. There is some freedom in choosing which bins to compare, in particular, in how to handle bins with zero charge. To be completely rigorous, one would compare every time bin of every DOM in the entire array for every event. Most of these, however, would have zero measured charge and nearly zero predicted charge, resulting in very expensively calculated multiplicative factor of one. This makes it attractive to drop zero-charge bins from the likelihood entirely, on the assumption that they will not contribute very much to the likelihood. This assumption is only valid at the true event parameters. Given the freedom to vary the event parameters, it is possible to maximize the likelihood by moving around in the parameter space until the worst-fitting parts of the predicted waveform are “hidden” in time intervals where the data-hypothesis mismatches do not contribute to the likelihood. Distortions of this sort could unfairly penalize the true event parameters and cause ΔLLH to become very large (see Section 5.6 and Figure 16).

There are additional assumptions, such as the basic correctness of the PHOTONICS photon propagation code and associated ice model, that cannot be tested in a closed simulation-reconstruction loop.

When this cascade reconstruction was developed, the added distortions of the full event simulation were accounted for by treating the simulation as a black box and calculating empirical correction factors to the PHOTONICS prediction [29, Section 5.4]. The total amplitude for an event, for example, was found on average to be only about 60% of the PHOTONICS prediction. The empirical correction factors proved quite useful in correcting a systematic underestimation of the energy. They were not successful in producing a self-consistent inversion of the simulation, as the discussion in Section 7.2 shows.

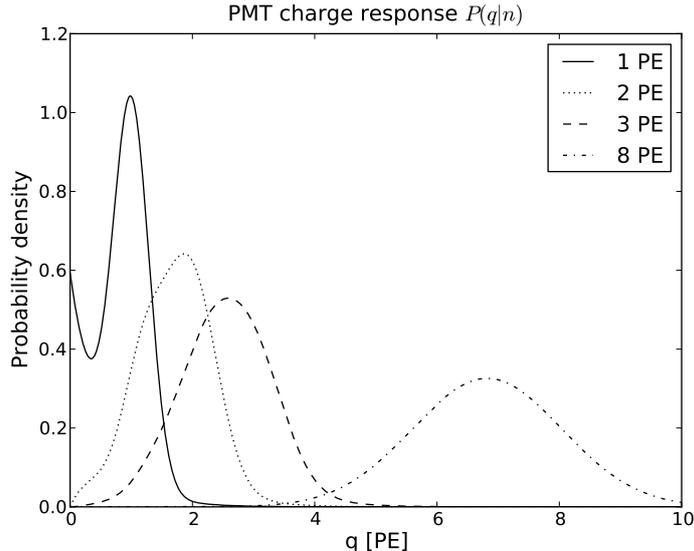


Figure 17: The charge response function $P(q|n)$ of the “standard” PMT used in ICECUBE simulation [4]. The response for $n > 1$ is obtained by repeatedly convoluting the single-photoelectron response function with itself. By 8 PE, it is very nearly gaussian.

The current study focuses on understanding the characteristics of a likelihood function that had already been shown to work in the sense that it produced an estimate of the event parameters close to the true event parameters. For this reason, it took a different approach to correction factors, breaking the simulation down to the simplest possible case and then building back up to a somewhat realistic simulation by introducing only correction factors whose meaning and effect could be understood and explained directly.

7.4 Deconstructing reconstruction

The simplest imaginable simulation is precisely the one modeled directly by the likelihood function: Poisson samples drawn from a distribution of means that is a function of the event parameters. This is equivalent to replacing the ICECUBE4 DOMs with ideal photoelectric detectors that can count individual photoelectrons with perfect accuracy and simulating only ideal, point-like Cherenkov emitters. When reduced to this level, the likelihood function should describe the fluctuations in the simulation.

The neutrino-propagation simulation produces a tree of energy-loss events that includes the final cascade in the detector. For the simplified simulation, the results of the propagation simulation were reduced to single ν_e undergoing charged-current interactions and depositing all their energy in point-like electromagnetic cascades in the ice³⁰. This guaranteed that the event actually matches a possible hypothesis used by the reconstruction. Photoelectrons (`I3MCHits`) were simulated from these events and used as the data for reconstruction with all correction factors removed. Even though all non-Poissonian elements had been stripped from the simulation, the ΔLLH distribution was similar to that of Figure 16. This indicated significant discrepancies between the PHOTONICS tables used for simulation and those used for reconstruction.

³⁰This is equivalent to modifying the physics such that only soft charged-current neutrino-nucleon scattering is allowed. In this contrived situation the 4-momentum transfer to the nucleon is negligible and there is no hadronic

When the tables were read out directly, the discrepancies were obvious. Figure 18 shows a comparison of the simulation and reconstruction table contents for a 100 TeV cascade 80 m from a DOM. One can gauge the severity of the discrepancies by estimating their contributions to the likelihood term for a single DOM. In each time bin, one can consider the log-likelihood of obtaining the value stored in the simulation table (λ_{sim}) from a Poisson distribution with a mean given by the value stored in the reconstruction table (λ_{reco}):

$$-\ln L = \ln \Gamma(\lambda_{\text{sim}} + 1) + \lambda_{\text{reco}} - \lambda_{\text{sim}} \ln \lambda_{\text{reco}} \quad . \quad (41)$$

This is a measure of the degree to which the likelihood penalizes the true event parameters. Figure 19 shows the log-likelihood contributions from each time-bin of the predicted PHOTONICS waveform. Since the amplitudes scale linearly with the cascade energy, the proportional discrepancies contribute more and more to the likelihood ratio (ΔLLH) for events with many photoelectrons, forcing the reconstruction away from the true vertex parameters.

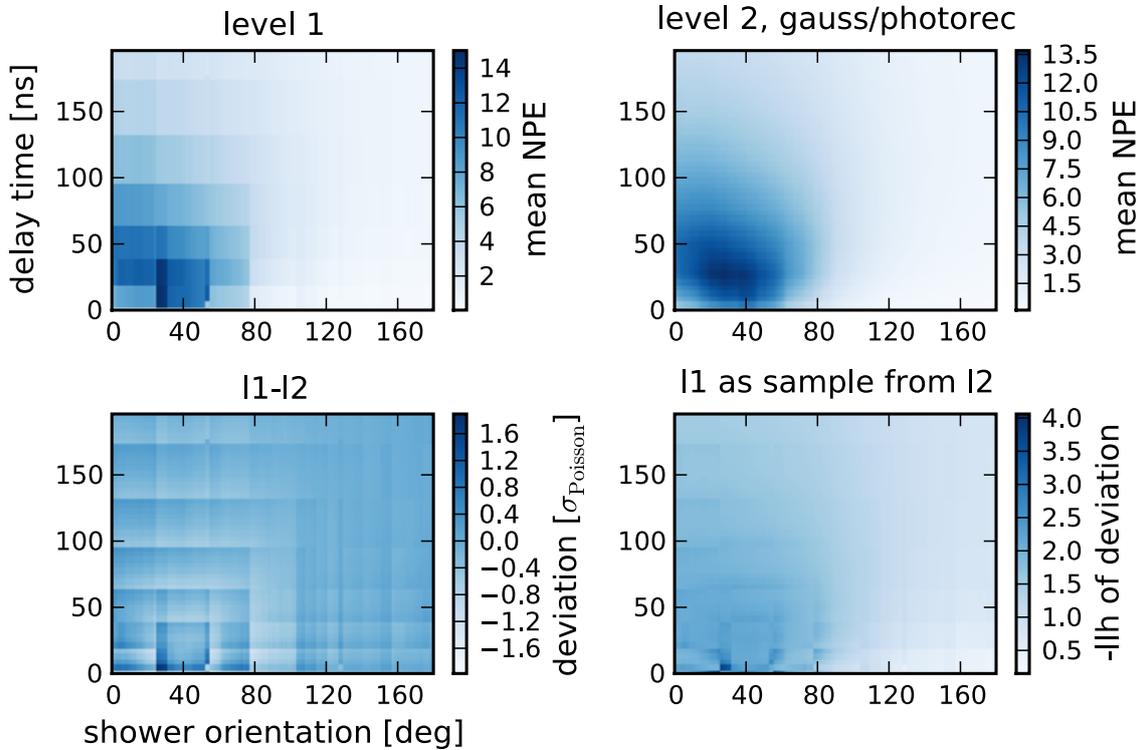


Figure 18: Comparison of Level 1 (simulation) and Level 2 (reconstruction) PHOTONICS predictions for a single DOM at depth $z = -310$ m. A 100 TeV cascade was placed at the same depth, 80 m away from and pointing towards the DOM, and rotated 180 degrees in azimuth. The two sets of tables were read out with the same interpolation settings. The upper left plot shows the Level 1 (simulation) mean amplitudes, while the upper right plot shows the same amplitudes for the Level 2 (reconstruction) tables. The differences in the interpolation schemes are clearly visible.

particle shower. Also, high-energy electromagnetic showers have no spatial extent; they are merely brighter.

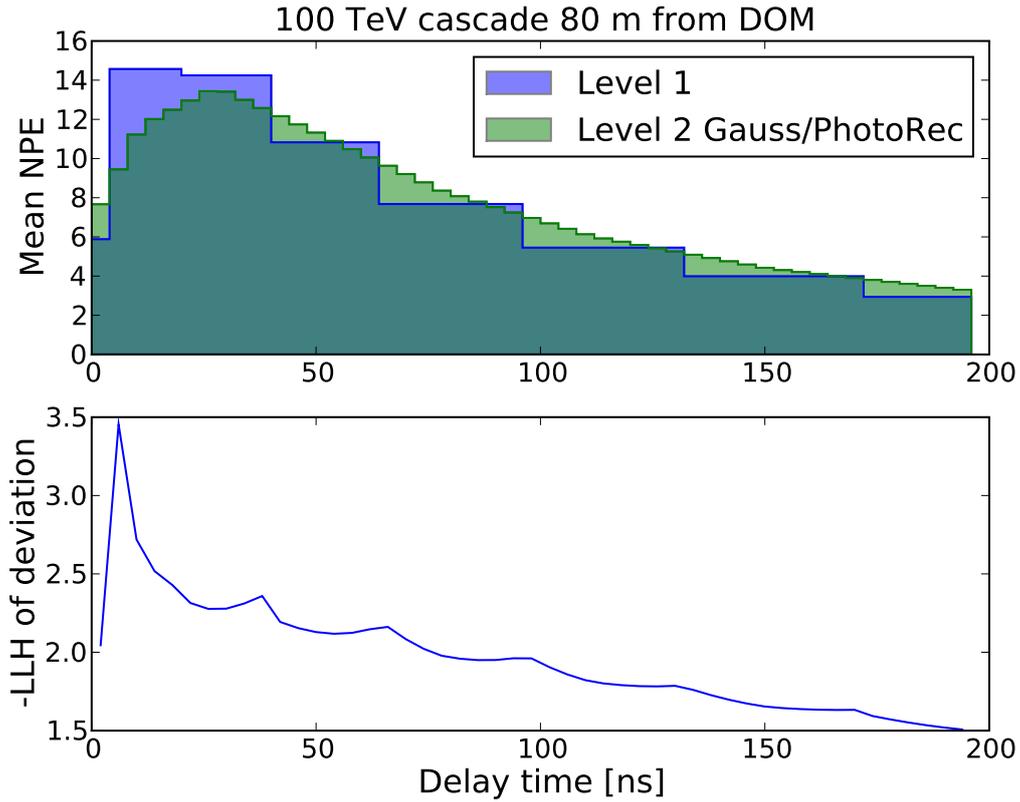


Figure 19: A slice of the distribution shown in Figure 18 at an azimuthal angle of 31 degrees. The two mean waveforms are sampled with the same granularity, but the Level 1 tables have no interpolation in time. The severity of the discrepancy can be estimated by calculating the negative logarithm of the probability of obtaining the Level 1 amplitudes as series of Poisson samples from the means given by the Level 2 amplitudes, as shown in the lower plot. The largest penalties are associated with the edges of the bins; events simulated at the edge of a bin in the Level 1 PHOTONICS tables will seldom reconstruct there.

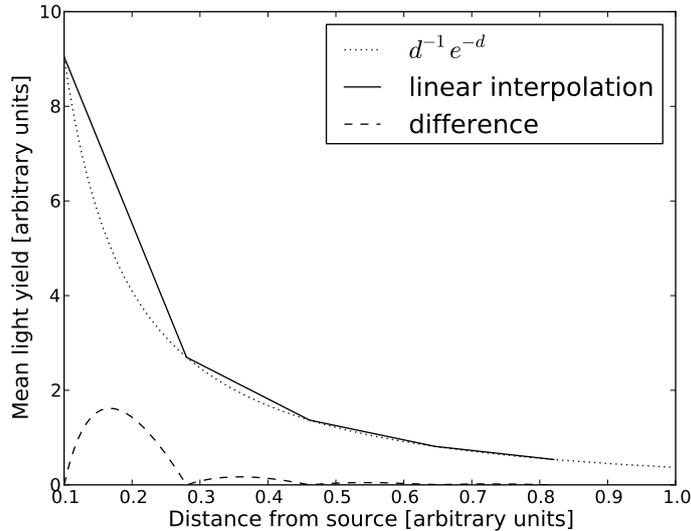


Figure 20: A schematic representation of how linear interpolation between bins in the PHOTONICS tables overestimates Cherenkov light detected in the ice.

7.5 Interpolation with B-splines

Having different linear interpolation schemes certainly introduces inconsistencies between simulation and reconstruction. These inconsistencies can only be detected if the reconstruction is sensitive enough to be affected by them. Beyond this, there are other problems caused by linear interpolation between PHOTONICS bins that are well-known. In particular, linear interpolation will always overestimate the total light yield³¹ as a function of distance from the source. The flux of Cherenkov photons in ice drops off exponentially as a function of distance from the source due to scattering and absorption [25, Equation 5.3.1]:

$$\mu \approx I_0 \frac{E}{d} e^{-d/\lambda_{\text{attn}}} \quad . \quad (42)$$

Figure 20 shows a schematic representation of how linear interpolation overestimates this light yield.

Large statistical fluctuations in underpopulated regions of the tables are also a concern. These effects can be mitigated by generating PHOTONICS tables with large numbers of tracked photons and with more closely-spaced bins, but this increases the size of the tables beyond the RAM capacity of most computers. Efforts were already underway to skirt the latter issue by instead fitting the fine-binned, high-statistics table to a multidimensional basis-spline (B-spline) surface with a coefficient array much smaller than the raw photon table [36]. This method, intended to produce a light distribution that more closely matches the physical one while smoothing out statistical fluctuations, also made it possible to use the same light distributions for simulation and reconstruction, removing a source of inconsistencies.

³¹“Light yield” and “photoelectrons ejected from the photocathode” can be used interchangeably, since each bin in the PHOTONICS table stores a flux of photons weighted to the angular acceptance of the DOM and frequency-dependent quantum efficiency of the photocathode.

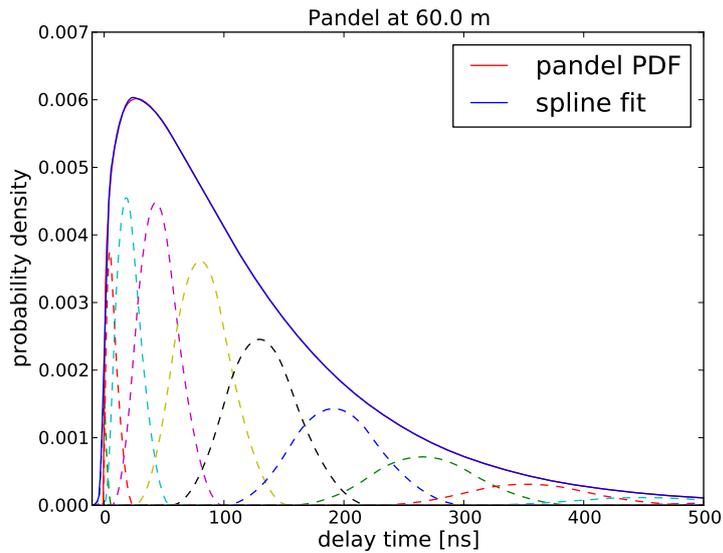


Figure 21: An example of a spline fit. Here, the photon delay time probability density function (PDF) calculated according to the model of [32] is represented in a 1-dimensional B-spline basis. The local gaussian like-functions (dashed lines) are the splines, and the points where two splines join are the knots. Each spline is identically zero outside of its support; for a second-order spline basis in 1 dimension, 3 splines have support at any given point. The 10 knots are spaced quadratically along the time axis. To evaluate the represented function at any point in time, one first evaluates the spline functions that have support at that point, then sums these evaluates weighted by the spline coefficients to obtain the function value. The resulting “spline fit” curve (the sum of the dashed curves) is able to follow the PDF much more closely than a piecewise linear function.

A rudimentary understanding of the spline technique is helpful for understanding why it is useful for creating alternative representations of the PHOTONICS tables. B-splines are local functions that consist of polynomial segments of low degree, usually quadratic or cubic. The points where splines meet are called knots. Since the splines have only local support, they can be used as a basis for modeling data with complex patterns [11], such as the differential photon distribution stored in each PHOTONICS table. The value of the differential distribution at each point in the 4-dimensional PHOTONICS coordinate space (ρ, l, ϕ, t) can be represented as a linear combination of the splines that have support at that point. The coefficients are determined by minimizing the squared differences between the spline surface and the data with an additional term that penalizes differences between the coefficients of neighboring splines, which smooths the resulting surface. This allows the surface to interpolate smoothly through regions of the table with large statistical fluctuations. Using polynomials of degree greater than one allows the spline surface to follow the curvature of the light distribution more closely, avoiding interpolation artifacts.

Since the fitting problem is linear in the spline coefficients, it can be represented as a deterministic matrix calculation. A naïve implementation of this computation for a data grid the size of a fine-binned PHOTONICS table would require excessive amounts of computer memory; a fast and compact version of this calculation is detailed in [11]. Once the coefficients have been determined, they are stored in a spline table along with the information necessary to evaluate the spline basis, such as the positions of the knots and orders of the splines along each dimension. The value of the spline surface at any point in the parameter space can then be determined by evaluating each spline weighted by its coefficient and summing over the splines that have support at that point. An ICETRAYS service was written to expose these functions to the ICETRAYS framework through the same interface as PHOTONICSSERVICE.

7.6 Simplified simulation and reconstruction with spline interpolation

In order to pinpoint the source of the reconstruction errors described in the previous section, the simulation was modified to use spline representations of the PHOTONICS tables to generate photoelectrons so that the same light distributions could be used in reconstruction.

7.6.1 Practical use of spline tables

A set of fine-binned, high-statistics tables was produced for the new simulation using the emission profile of an electromagnetic cascade. Since the Antarctic ice is not an isotropic medium, the results of the PHOTONICS simulation depend on the depth and zenith angle of the source. Generating a set of tables for the full detector volume takes an extraordinarily large amount of CPU time³², so for this proof-of-concept study the ice properties at one depth were taken for the entire detector. The measured optical properties at this depth³³ are close to properties averaged over the entire detector volume [15], so this “bulk ice” provides a reasonable starting point for tests of simulation and reconstruction. The binning schemes of the “standard” and “new” tables are given in Table 1.

³²Simulating the propagation of 4×10^7 photons through the ice layers for all combinations of 158 depth bins and 18 zenith angle bins takes on the order of 19000 hours on currently available hardware.

³³This depth is between two dust layers in the upper half of the detector, 1720 m below the surface at ICECUBE z -coordinate 226 m.

Coord	Standard “AHA v1”			New tables			Bin spacing
	Limits		# Bins	Limits		# Bins	
ρ [m]	0	580	30	0	580	200	quadratic
ϕ [deg]	0	180	9	0	180	36	equal
l [m]	-580	580	65	-580	580	100	quadratic
t [ns]	0	7000	50	0	7000	105	quadratic
θ [deg]	0	180	18	0	180	18	equal
z [m]	-815.4	764.6	158	226	226	1	equal

Table 1: A comparison of the standard PHOTONICS tables used for cascade simulation and the new tables generated for this study. The former simulation propagated 4×10^7 photons; the one for this study used 8.9×10^8 to reduce statistical fluctuations associated with the finer binning.

Each of the generated tables³⁴ was differentiated along the time axis and its logarithm fit with a 4-dimensional B-spline surface using the algebraic least-squares algorithm described in [11]. The logarithm of the differential was taken to ensure the monotonicity of the cumulative distribution. The set of spline surfaces was then merged using a heuristic to interpolate smoothly between adjacent spline surfaces representing different zenith angles of the source. The total amplitude tables³⁵ were also fit to 3-dimensional spline surfaces and merged in the same way.

The interface to PHOTONICSSERVICE allowed these spline tables to be used with the existing simulation and reconstruction software with minimal changes. One minor change that was necessary was due to the fact that the spline tables only store the differential distribution. The standard PHOTONICS tables used for simulation store the cumulative light distribution in delay time, $F(t) \in [0, 1]$. This distribution can easily be sampled by generating random numbers $q \in [0, 1]$ and inverting for $t = F^{-1}(q)$ by binary search. This is efficient if the expected photoelectron yield is smaller than the granularity of the simulated electronics. For bright events, however, it becomes more efficient to operate in “binned mode,” where the differential distribution is sampled by integrating over a bin by the midpoint method and drawing a random number from a Poisson distribution with a mean given by the mean amplitude obtained from the bin integral. Since only the differential distribution was available, the simulation was modified to use only binned mode to generate hits.

This change created a few pathologies of its own. First, because the fitting algorithm is not currently able to account for restrictions on the integral of the spline surface along a particular dimension, the spline tables are not necessarily normalized. This means that the integral of the timing table is not necessarily that same as the value stored in the amplitude table. This makes it difficult to reliably predict the total charge in a DOM, since the total charge was simulated by integrating over the timing table. Secondly, the approximation used to integrate the waveform depends strongly on the point at which the differential is calculated from the table, which in turn depends on the time binning used to simulate the electronics. At points close to the source where unscattered photons create delta-like time distributions, the time-derivative of the differential distribution can be very large, and moving the center of the first time bin by only a nanosecond can change the value of the integral by factors of 2 or more. This is less of a problem, since the approximation of a point-like cascade cannot be expected to hold at distances where there is no significant scattering. A work-around was added to catch cases where the calculated mean in

³⁴Each table contains bins in ρ , ϕ , l , and t and represents a single depth and polar orientation of the source. There were 18 tables in this set, each for a particular orientation of the source.

³⁵The amplitude tables produced by PHOTONICS have the time dimension integrated out, so that they contain only the total expected light yield at each point.

one bin was larger than the integral stored in the amplitude table. When such a bin was found, the amplitude in the bin was set to zero so as not to contribute to the likelihood value.

7.6.2 Baseline reconstruction

With these corrections in place, the spline tables were used to generate hits for point-like cascades; these hits were used as the basis for reconstruction as described in Section 7.4. The residual as a function of photoelectrons (NPE³⁶) is shown in Figure 22. The bulk residual scales with NPE across the entire range of brightnesses studied, indicating that the reconstruction error is dominated by statistics in this case rather than systematics. This is expected, since a major source of systematic error was removed when a consistent PHOTONICS table interpolation scheme was introduced.

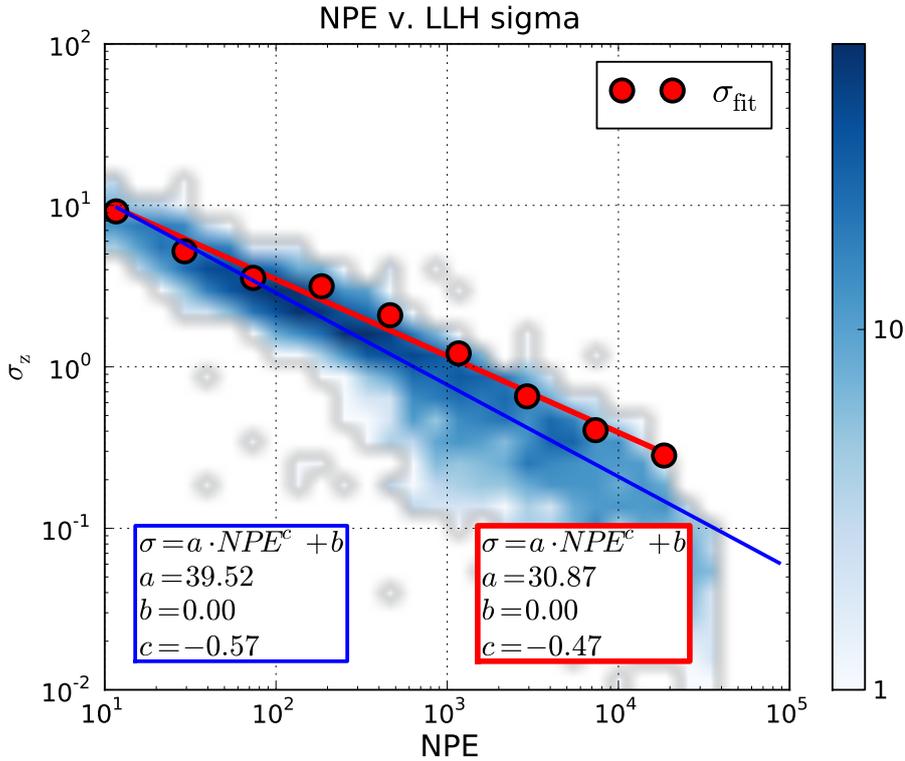


Figure 22: Resolution in vertex depth z (σ_z) as a function of total number of photoelectrons (NPE) for point-like cascades simulated with perfect electronics in bulk ice. The slices of the bulk residual distribution follow the variance of the likelihood function much more closely, but the variance of the likelihood function still underestimates the residual. The bulk residuals scales with NPE across the entire range of NPE, indicating that the reconstruction is now dominated by statistics rather than systematics.

The ΔLLH distribution in Figure 23 reveals additional systematic errors that were not yet accounted for in the simulation. In particular, the exponential scaling of ΔLLH with NPE

³⁶Since the hadronic cascades in the sample were replaced with point-like electromagnetic cascades, the relationship between total charge and cascade energy is different from the standard simulation. This is shown in Figure 35.

is indicative of a proportional offset in the expected light yield that becomes more and more significant at higher intensities, where the expected relative fluctuations are smaller.

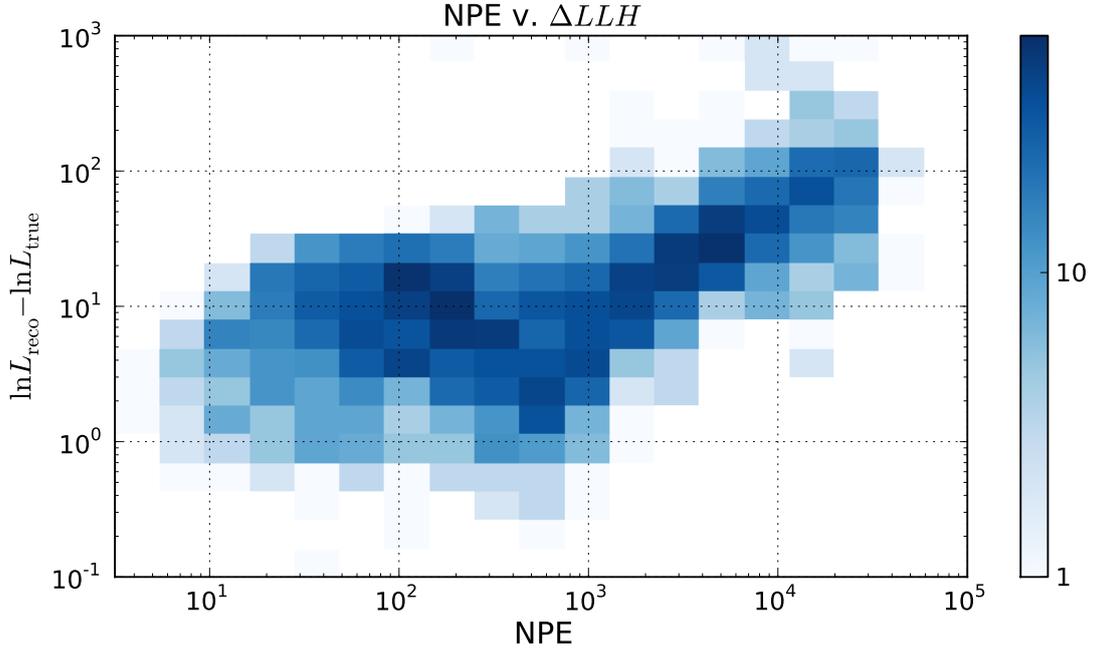


Figure 23: Log-likelihood ratio as function of total number of photoelectrons (NPE) for point-like cascades simulated with perfect electronics in bulk ice. The ratio is well-behaved for events with 500 PE, but becomes excessively large for events with higher and lower photoelectron counts, indicating that the likelihood formulation still does not accurately describe the simulation.

7.6.3 Total amplitude correction

The scaling of ΔLLH at high NPE was discovered to be due to an overlooked correction factor in the hit-generation code. By default, the hit-generation code scales the PHOTONICS mean amplitude down by 10% to account for the fraction of the photocathode area that is shadowed by the string cable. Since the reconstruction did not contain this correction factor, the likelihood was distorted whenever the expected fluctuations became small compared to 10% of the total amplitude. Adding this correction factor brought the likelihood ratio under control at high NPE as shown in Figure 24. The remaining cases of large ΔLLH are concentrated at low NPE, which seems strange at first, since those events should also have the largest allowed fluctuations.

7.6.4 Zero-charge bins

These remaining cases were an artifact of another approximation made in the formulation of the likelihood. In an effort to improve performance, the likelihood function did not consider contributions for zero-charge (or un-hit) bins. This selective binning has a significant effect for sparsely-populated DOMs; a higher total amplitude is a much better fit when the un-hit bins are not considered at all. In the original formulation of the likelihood, these bins were lumped together in a term that compared the integrated charge in the DOM with the total expected charge. This was dropped from the spline formulation, as normalization mismatches between

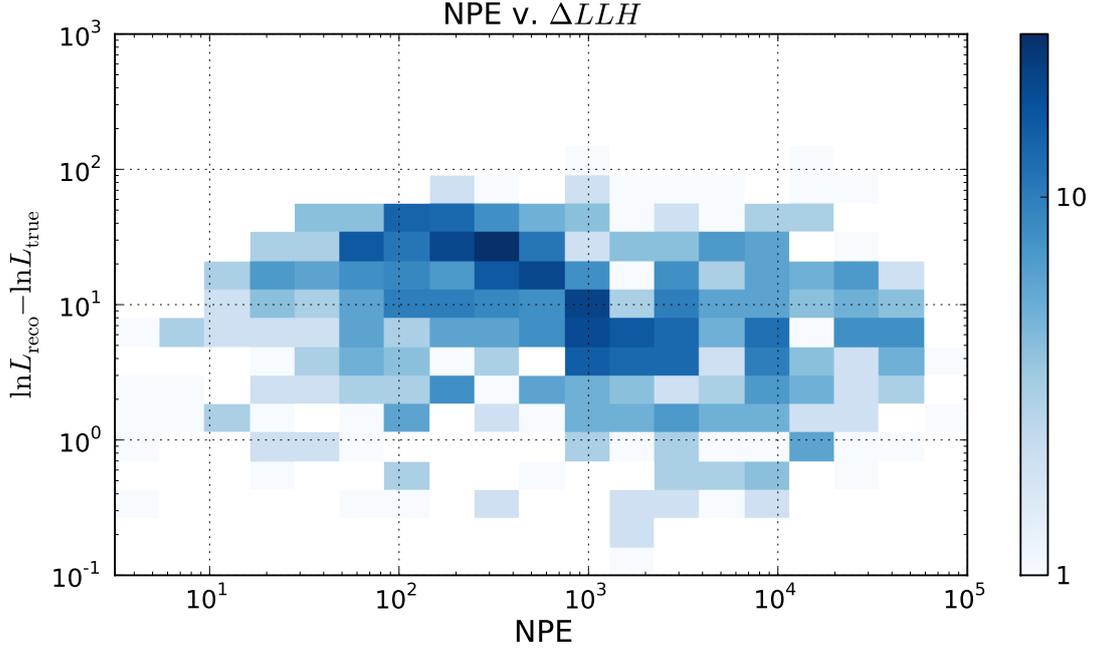


Figure 24: Log-likelihood ratio as function of total number of photoelectrons (NPE) for point-like cascades simulated with perfect electronics in bulk ice, with a correction factor applied to account for a global offset in the total PE yield. While this was the result of a mistake, it illustrates the effect of a constant systematic offset on a Poisson likelihood.

the timing and amplitude tables would result in constant proportional offsets for some points in the tables, leading to effects similar to those shown in the high-NPE portion of Figure 23. Without this constraint, the correct solution to the inference problem is highly suppressed in favor of a false solution. For brighter and brighter events, the average number of un-hit bins per DOM approaches zero, and the approximation is no longer used. To work around this issue without making the calculation excessively slow, un-hit bins were grouped together and considered as a single, wider bin. While this still treats hit and un-hit bins differently³⁷, it is sufficient to constrain the total predicted amplitude to reasonable values. When the un-hit bins are considered in the likelihood calculation, the true event parameters are finally statistically compatible with the reconstructed values. This is true across the entire range of energies in the sample as Figure 25 shows.

7.7 Reconstruction with electronics simulation

7.7.1 Correcting for the PMT charge response

Once the likelihood had been made to behave well with hits only, it was possible to add the PMT and readout electronics back in to the simulation. Instead of hits, the simulation then produced structures of ADC counts which had to be converted to voltages using each DOM's calibration constants. This series of voltages is a convolution of the waveform on the secondary winding of the PMT base transformer with the amplifier pulse shapes. The PMT waveform can be deconvoluted from the measured waveform using the Bayesian unfolding algorithm implemented

³⁷In particular, the midpoint integral approximation can be problematic for long intervals of un-hit bins.

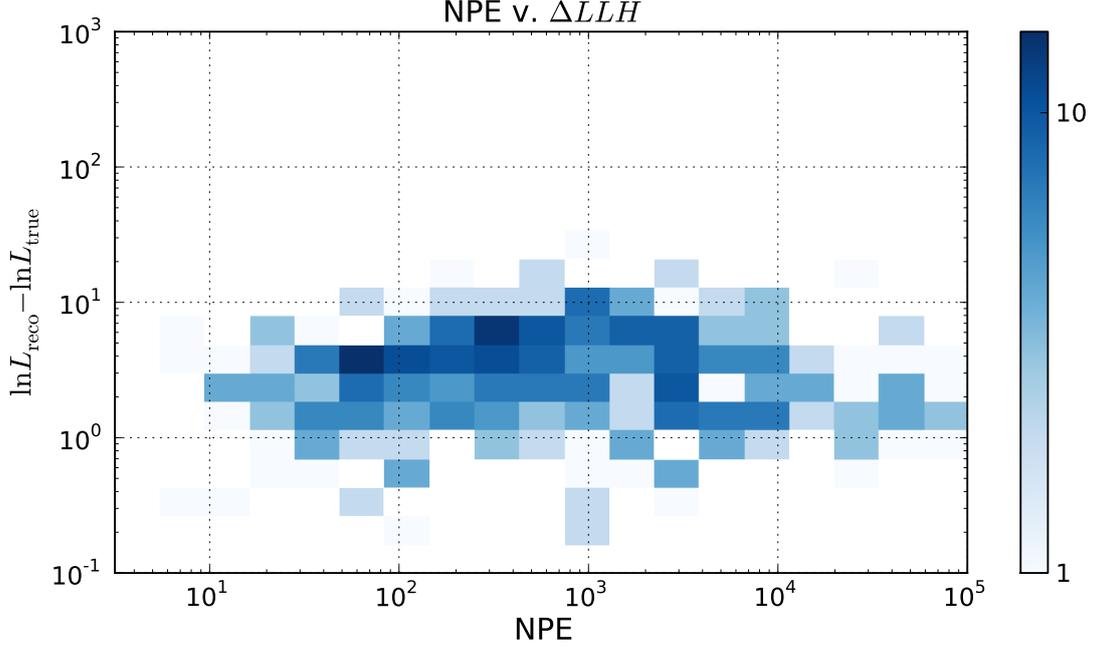


Figure 25: Log-likelihood ratio as function of total number of photoelectrons (NPE) for point-like cascades simulated with perfect electronics in bulk ice, with terms included in the likelihood for un-hit bins. The likelihood ratio is now independent of NPE, as expected.

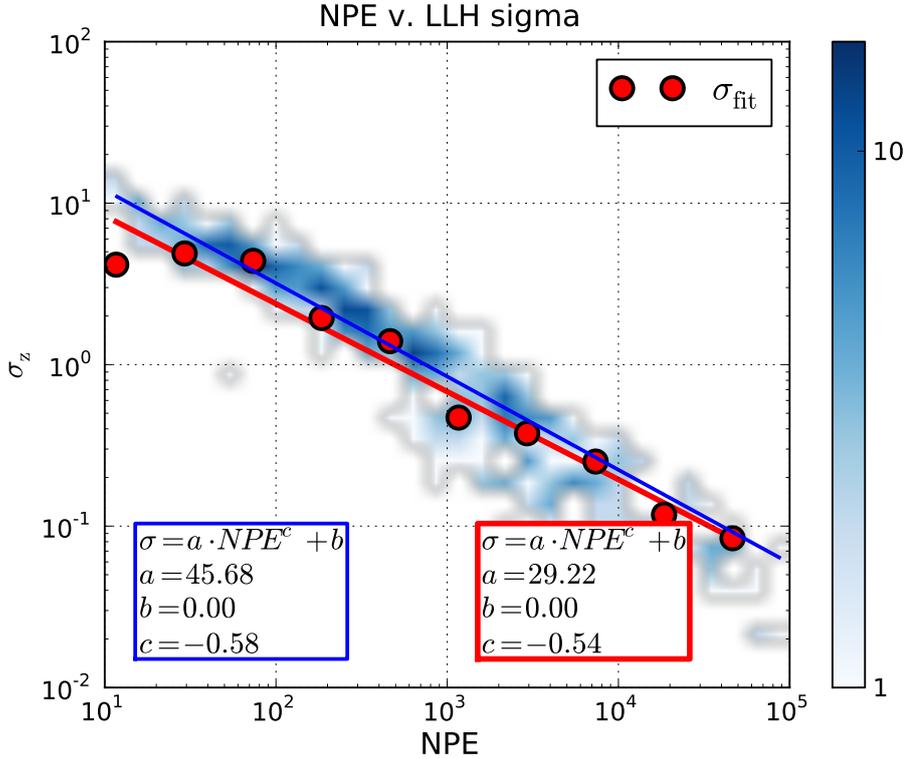


Figure 26: Resolution in vertex depth z (σ_z) as a function of total number of photoelectrons (NPE) for point-like cascades simulated with perfect electronics in bulk ice. The variance of the likelihood function is now a good estimate of the residual.

in PULSEEXTRACTOR [5] to produce reconstructed pulses, each of which represent the best guess at the integrated photocathode current in a particular time window.

Unlike hits, pulses no longer come in integer units, as they are obtained from the amplified anode current, and thus also include fluctuations due to the $\sim 30\%$ charge resolution of the PMT. The additional fluctuations can be accounted for exactly by summing the charge response contributions $P(q|n)$ (see Figure 17) from each possible number of photoelectrons:

$$P(q|\lambda) = \sum_{n=0}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} P(q|n) \quad (43)$$

The infinite sum here could become quite expensive, especially since the PMT charge response function for > 1 PE is itself obtained by numerical convolution. In practice, however, it was sufficient to replace the factorial in the Poisson probability with a gamma function to smoothly interpolate between integer units of charge:

$$P(q|\lambda) \approx \tilde{P}(q|\bar{q}_1\lambda) = \frac{e^{-\bar{q}_1\lambda} (\bar{q}_1\lambda)^q}{\Gamma(q+1)} \quad (44)$$

where \bar{q}_1 is the mean of the single-photoelectron charge response function for a particular DOM (0.850446 for the standard PMT used in simulation) and λ is the PHOTONICS prediction scaled by the cable shadow correction factor of 0.9. Figure 27 shows a comparison of the approximate probability distribution of Equation (44) with the more realistic PMT base current distribution of Equation (43). There is a fairly large effect when the expected photocathode current is small; the analytic approximation overestimates the probability of obtaining fractional charges. At higher mean amplitudes, the analytic approximation is slightly too sharply peaked, since it doesn't account for contributions from the tails of the charge response function. Despite this, the approximation provides a reasonable estimate of the characteristic scale of fluctuations in the PMT base current, making it possible to use the full likelihood function to estimate the statistical error of a reconstruction.

With the electronics included in the simulation, the reconstruction also has to account for PMT saturation. When a small number of photoelectrons are ejected from the PMT cathode, the mean base current is proportional to the photocathode current. The constant of proportionality is the nominal gain, typically on the order of 10^7 . For large photocathode currents, however, the response can be less than linear. The saturation behavior of the PMT can be measured and parameterized as a function of the instantaneous PMT illumination (see e.g. Figure 10 of [35]), but is slightly different for each PMT. A typical IceCube PMT will show a 10% non-linearity at a photocathode current of 31 PE/ns [35]. The current simulation does not account for differences in the saturation behavior of the PMTs in different DOMs, and thus implements a single saturation curve. There is almost an order of magnitude between the current at which saturation becomes significant and complete saturation, that is, a sizable region in which the saturation function is invertible. In principle, it is possible to correct the charge of the reconstructed pulses using the inverse of the saturation function to recover some of the dynamic range of the PMT. Since individual saturation curves were not available, however, it was more expedient to simply exclude the saturated DOMs from the likelihood calculation.

In the original implementation of the reconstruction, a DOM was classified as saturated or linear based on its total charge. In reality, however, the saturation behavior of long and short pulses is nearly identical [35, Figure 10], so saturation is just a function of instantaneous photocathode

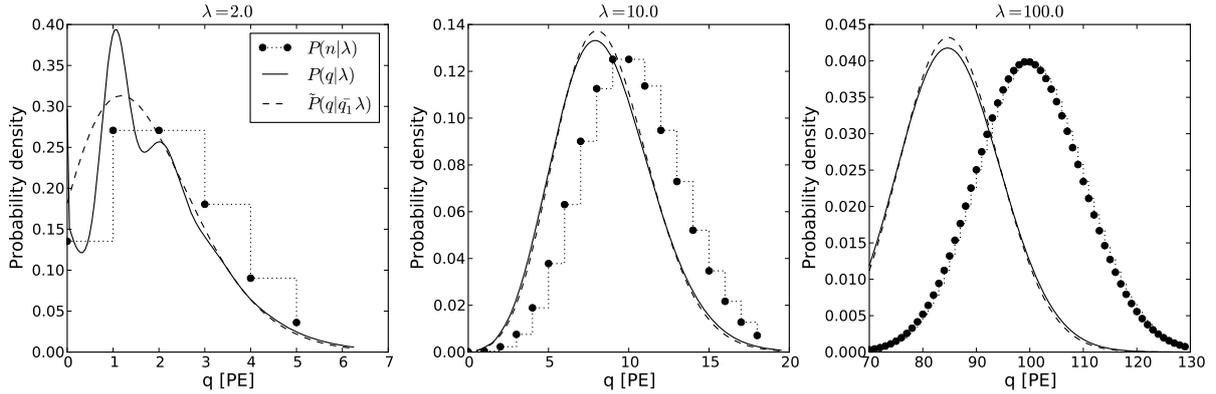


Figure 27: An approximation to the distribution of charge responses given the photoelectron yield at the photocathode. The stepped line shows the Poisson probability of ejecting n photoelectrons from the photocathode. The solid line shows the numerical convolution of the photocathode current distribution with the charge response function (see Figure 17) as given in Equation (43), and the dashed line shows the analytic approximation using the extension of the Poisson probability distribution to real numbers q given in Equation (44). The mean of the continuous-Poisson distribution was multiplied by \bar{q}_1 (~ 0.85) to account for the reduced efficiency of the PMT.

current³⁸, and the total charge collected by a DOM is a poor indicator of the presence of saturated pulses. The way to find saturated DOMs, then, is to look for pulses representing a photocathode current in the saturation region. The appropriate cutoff can be estimated by plotting the saturation curve against the variance of a Poisson distribution as shown in Figure 28. The saturation level for this simulation was chosen as 50 PE/ns, since for larger currents the Poisson fluctuations are small compared to saturation effects. DOMs with pulses representing average currents of more than 50 PE/ns were excluded from the likelihood calculation entirely rather than just excising the saturated pulses to avoid distortions due to gaps in the waveform³⁹.

7.7.2 Reconstruction in restricted dimensions

With these corrections in place, it was possible to reconstruct point-like electromagnetic cascades with fixed energy and direction using pulses obtained from the output of the electronics simulation. Figure 29 shows that the ΔLLH distribution is still independent of the total event charge, indicating that the electronics simulation does not introduce significant systematic errors in this configuration. The variance of the likelihood function is a good indicator of the residual as shown in Figure 30.

7.7.3 Reconstruction in all dimensions

Up to this point, all reconstructions had been done with the cascade orientation and energy fixed at the true values. Once the reconstruction had been made to work in this slice of the likelihood space, it was possible to allow the remaining parameters to vary freely as well. Allowing the

³⁸This behavior is also the one modelled in the simulation.

³⁹The effect of gaps in the waveform should be similar to the distortions caused when un-hit bins were excluded from the likelihood calculation, shown in the low-brightness side of Figure 24.

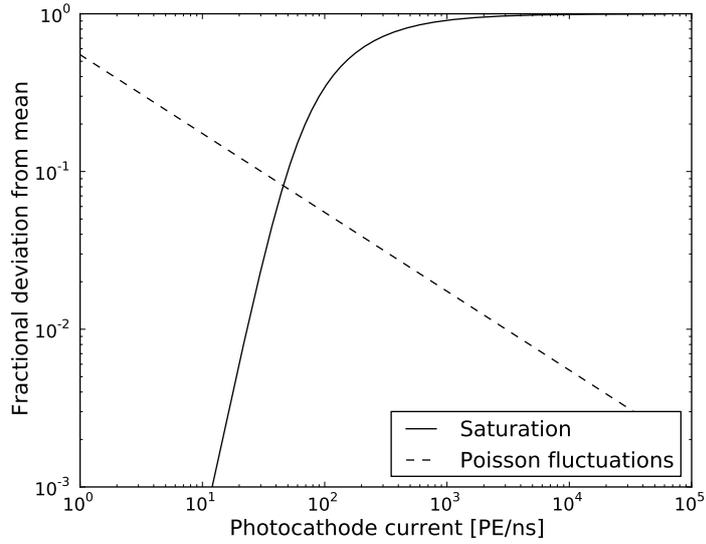


Figure 28: A comparison of deviations from the expected mean PMT base current due to Poisson fluctuations in the photocathode current and non-linearity of the PMT response, using the saturation curve defined in [4]. The Poisson fluctuations are taken as $1/\sqrt{N}$ fluctuations in the integrated current over a time window of 3.3 ns, approximately the sampling time of the ATWD. For the standard PMT used in ICECUBE simulation, saturation effects are larger than the expected fluctuations in photocathode current at photocathode currents larger than 50 PE/ns.

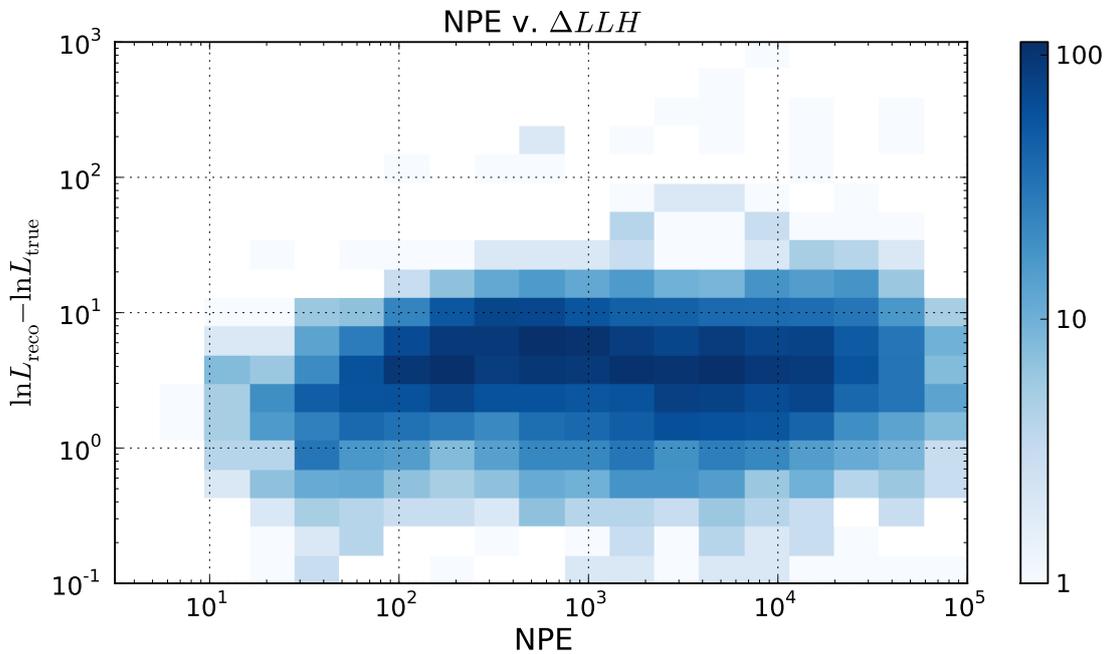


Figure 29: Log-likelihood ratio as function of total event charge (NPE) for point-like cascades simulated with realistic electronics in bulk ice. The likelihood ratio is now independent of NPE.

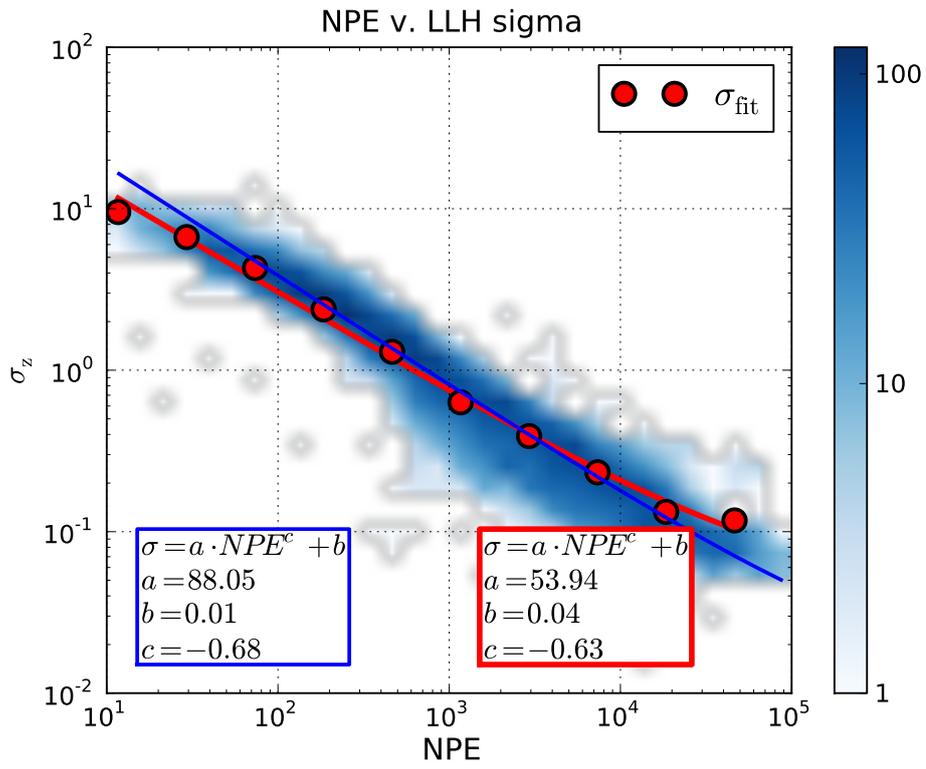


Figure 30: Resolution in vertex depth z (σ_z) as a function of total event charge (NPE) for point-like cascades simulated with realistic electronics in bulk ice.

cascade energy to vary freely in particular introduces an important additional degree of freedom to the problem. Whereas with fixed energy the likelihood function attempts to fit the observed light distribution to the one parameterized by the PHOTONICS tables, allowing the energy to vary allows the overall normalization to vary as well. This makes it possible to obtain a better solution by moving the vertex far away from the detector and scaling the energy up to match the observed light yield. This problem is exacerbated by the omission of un-hit DOMs from the likelihood, since the absence of charge in some DOMs places a strong constraint on the overall light yield.

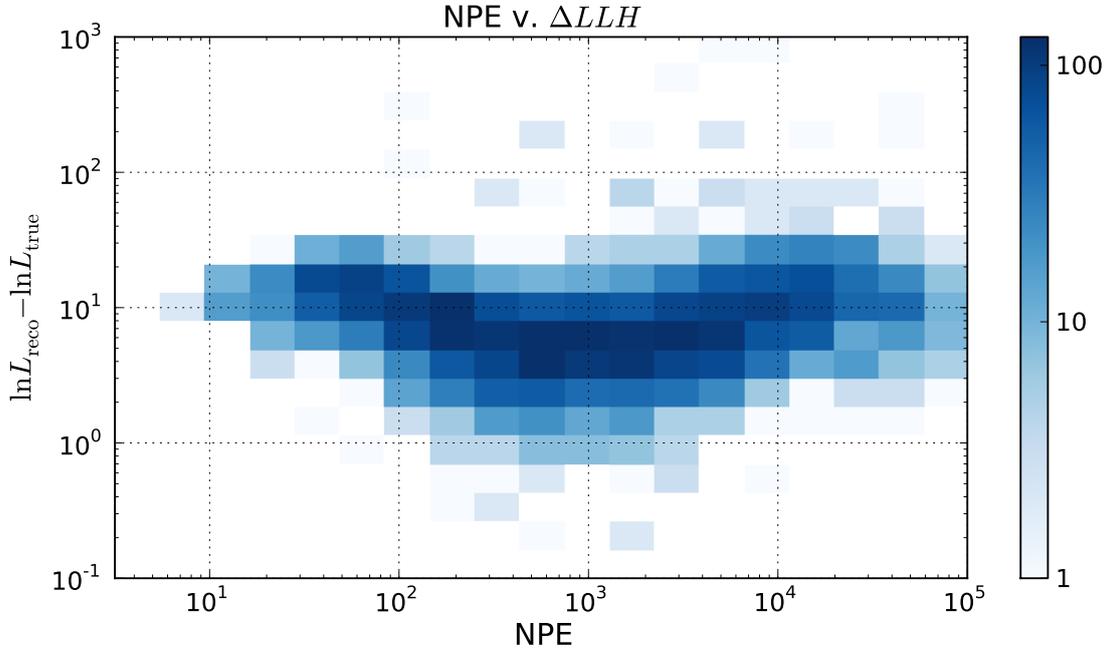


Figure 31: Log-likelihood ratio as function of total event charge (NPE) for point-like cascades simulated with realistic electronics in bulk ice when all 7 event parameters were allowed to vary freely.

When un-hit DOMs are included in the likelihood calculation⁴⁰, the solution is much more tightly constrained for low-brightness events, as shown in Figure 33. Plots of the resolution in zenith angle and cascade energy in this scenario are given in Appendix B.

⁴⁰It is important to continue to exclude broken DOMs from the calculation, as these would distort the likelihood function.

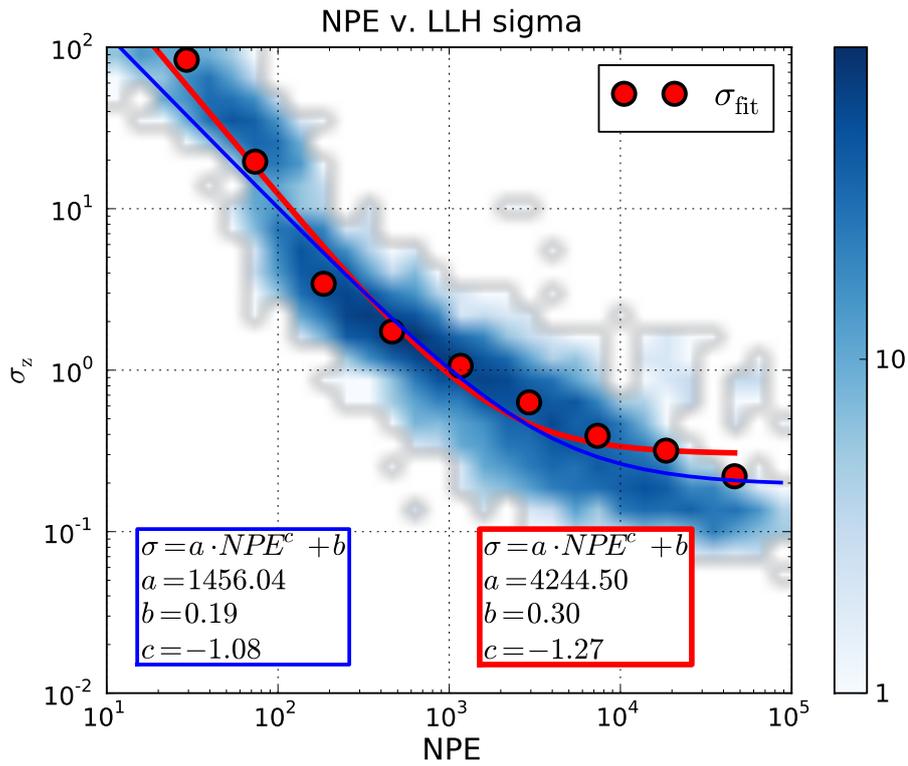


Figure 32: Resolution in vertex depth z (σ_z) as a function of total event charge (NPE) for point-like cascades simulated with realistic electronics in bulk ice when all 7 event parameters were allowed to vary freely.

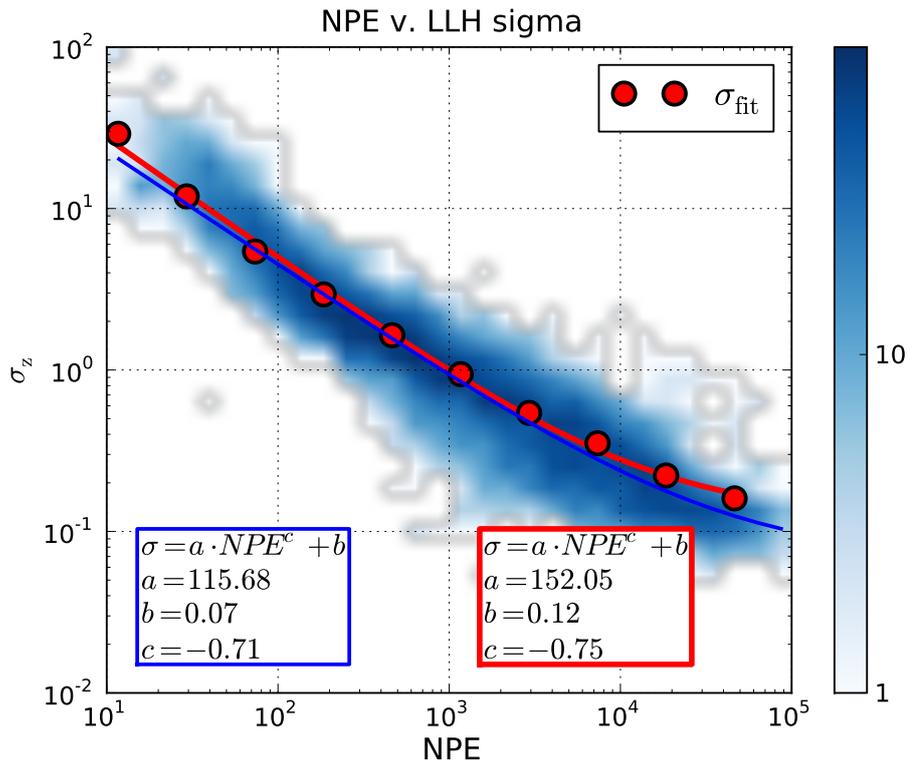


Figure 33: Resolution in vertex depth z (σ_z) as a function of total event charge (NPE) for point-like cascades simulated with realistic electronics in bulk ice, with un-hit DOMs included in the likelihood calculation.

8 Conclusion

This work has demonstrated how Markov-Chain Monte Carlo techniques can be used to estimate the inherent resolution of a likelihood-based reconstruction. Furthermore, it has shown that a better statistical model can lead to both smaller residuals and more accurate estimates of those residuals. In particular, the systematic component of the resolution improved significantly, as shown in Tables 2 and 3.

Scenario		x	z	time	zenith	energy
Standard simulation	a	1.50×10^2	8.62×10^1	2.08×10^3		
	b	1.26×10^1	8.55	4.57×10^1		
Minimal simulation with splines	a	4.33×10^1	3.09×10^1	1.80×10^2		
	b	0.73	1.31×10^{-8}	1.89		
+ 10% amplitude correction	a	6.13×10^1	4.10×10^1	1.48×10^2		
	b	5.41×10^{-10}	1.87×10^{-9}	5.35×10^{-8}		
+ no-hit probability	a	3.30×10^1	2.92×10^1	8.12×10^1		
	b	4.37×10^{-2}	5.07×10^{-9}	1.30×10^{-8}		
+ electronics	a	2.31×10^1	5.39×10^1	9.27×10^1		
	b	0.26	3.91×10^{-2}	1.82×10^{-8}		
+ all dimensions	a	1.20×10^3	4.24×10^3	9.40×10^2	5.49	1.32×10^2
	b	1.07	0.30	1.51×10^{-7}	1.26×10^{-2}	8.43×10^{-3}
+ un-hit DOMs	a	8.27×10^1	1.52×10^2	2.46×10^3	4.37	3.60
	b	0.51	0.12	1.45	6.05×10^{-3}	6.14×10^{-3}

Table 2: Scaling behavior of the residual distribution with total event charge (NPE). The parameter a gives the coefficient of the statistical component of the resolution, which dominates at low intensities, while b gives the systematic error floor, which dominates at high intensities. These are the same parameters shown for example in the fit in Figure 15.

It is important to note that the simplified simulation used in this demonstration does not reflect all the effects known or thought to influence real ICECUBE data. Section 7.3 gave a list of implicit approximations in the statistical model represented by the likelihood function. The simplified simulation used in this demonstration removed some components of the simulation chain in order to make these approximations exact. Some of these component were reintroduced with along with proper correction factors in the likelihood function.

The following effects were responsible for significant distortions to the likelihood function.

Interpolation artifacts in Photonics tables By far the largest improvement comes from the use of identical light yield distributions in simulation and reconstruction. The fact that they are consistent with each other does not imply that they are consistent with reality. Comparisons with direct photon tracking and flasher data should be done.

PMT transit time and pulse shape The inherent timing resolution of the PMT smears out the photoelectron arrival times, and this must be properly accounted for by convoluting

Scenario		x	z	time	zenith	energy
Standard simulation	a	4.53×10^1	4.61×10^1	2.13×10^2		
	b	8.64×10^{-9}	8.69×10^{-2}	0.40		
Minimal simulation with splines	a	2.23×10^1	3.95×10^1	8.12×10^1		
	b	3.41×10^{-9}	2.03×10^{-8}	8.69×10^{-8}		
+ 10% amplitude correction	a	1.46×10^1	3.82×10^1	1.08×10^2		
	b	4.86×10^{-9}	1.71×10^{-8}	1.81×10^{-9}		
+ no-hit probability	a	4.90×10^1	4.57×10^1	9.33×10^1		
	b	2.40×10^{-8}	6.96×10^{-10}	1.10×10^{-8}		
+ electronics	a	5.96×10^1	8.81×10^1	9.49×10^1		
	b	0.12	1.17×10^{-2}	7.38×10^{-8}		
+ all dimensions	a	2.18×10^3	1.46×10^3	1.77×10^6	1.13×10^1	5.97×10^2
	b	0.61	0.19	1.88	4.25×10^{-3}	9.06×10^{-3}
+ un-hit DOMs	a	1.93×10^2	1.16×10^2	1.53×10^3	7.62	1.18
	b	0.26	6.70×10^{-2}	0.78	5.14×10^{-3}	2.64×10^{-3}

Table 3: Scaling behavior of the likelihood distribution with total event charge (NPE). The parameter a gives the coefficient of the statistical component of the resolution, which dominates at low intensities, while b gives the systematic error floor, which dominates at high intensities.

the expected waveform with the expected distribution of leading-edge times of the pulses reconstructed from digitized data. In this simplified simulation, the distribution of reconstructed arrival times was adequately described by a normal distribution with $\sigma = 2.18$ ns, the same width as the simulated PMT pulse.

Saturation Non-linear PMT response to large photocathode currents can distort the reconstruction of bright events. Simply excluding the affected DOMs from the reconstruction gets rid of the distortion, but also loses some of the most useful information, since the DOMs closest to the event are first ones to saturate. It would be better to either invert the saturation curve or account for the distortion in the likelihood function, but this would also require that the charge response of each DOM be simulated individually in order to be useful.

Zero-charge bins Un-hit bins in the waveform turned out to be an important constraint on the event parameters; excluding the terms arising from zero-charge bins from the likelihood function unfairly penalized the true event parameters in favor of alternative solutions. Grouping contiguous zero-charge bins together seems to provide an adequate compromise between accuracy and performance.

Other effects didn't seem to distort the reconstruction in any significant way.

PMT charge response The charge resolution of the PMT can be adequately accounted for by extending the Poisson distribution to non-negative real numbers and scaling down the mean, as shown in Figure 27.

Waveform reconstruction The Bayesian unfolding algorithm implemented in PULSEEXTRACTOR accurately recovers the arrival photon arrival time distribution to within the width of the simulated PMT pulse and preserves the total charge of the waveform by construction. Given an adequate representation of the amplifier pulse shape, the pulse reconstruction does not distort the reconstruction significantly.

There remain a few effects that were not adequately investigated in this work.

Light emission profile Real electromagnetic and hadronic particle showers are not point-like. In the production simulation, the longitudinal extent of cascades is approximated by distributing sub-cascades along the direction of the track according to an energy deposition distribution. The simplified simulation used in this work replaced the spatially-extended cascade with a point-like cascade of the same energy. If accounted for properly in the reconstruction, the longitudinal extent of the cascade would add a geometrical constraint to the reconstruction, which should improve the directional resolution at energies.

Spurious PMT pulses Since spurious PMT pulses are much rarer than “normal” pulses, they tend to significantly distort the expected waveform only when the expected waveform has a sharp, high-amplitude peak. Since waveforms in this category are also likely to be saturated, the tests in this work were unlikely to be sensitive to the effects of spurious pulses.

Once these effects have been accounted for in the formulation of the likelihood function, the techniques presented in this work could be extended to a full detector simulation with a layered ice model, and from there to experimental data.

References

- [1] IceRec doxygen documentation. Available from: <http://www.ifh.de/~kislak/www/icerec/docs/V03-00-00/doxygen/>.
- [2] IceSim doxygen documentation. Available from: <http://www.physto.se/~klas/icecube/icesim-doxy/>.
- [3] IceTray public Subversion repository. Available from: <http://code.icecube.wisc.edu/icetray-dist/>.
- [4] PMTSimulator source code. Available from: <http://code.icecube.wisc.edu/svn/projects/pmt-simulator/trunk/>.
- [5] PulseExtractor source code. Available from: <http://code.icecube.wisc.edu/svn/sandbox/PulseExtractor/>.
- [6] George E. P. Box and Gwilym M. Jenkins. *Time series analysis: forecasting and control*. Holden-Day, San Francisco, 1970.
- [7] Siegmund Brandt. *Statistical and Computational Methods in Data Analysis*. North-Holland Publishing Company, Amsterdam, 1970.
- [8] Giulio D’Agostini. A Multidimensional unfolding method based on Bayes’ theorem. *Nucl. Instrum. Meth.*, A362:487–498, 1995. doi:10.1016/0168-9002(95)00274-X.
- [9] Luke Drury. An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas. *Rep. Prog. Phys.*, 46:937–1027, 1983. doi:10.1088/0034-4885/46/8/002.
- [10] Anthony F. W. Edwards. *Likelihood: an account of the statistical concept of likelihood and its application to scientific inference*. Cambridge University Press, 1972.
- [11] Paul H. C. Eilers, Iain D. Currie, and María Durbán. Fast and compact smoothing on large multidimensional grids. *Computational Statistics and Data Analysis*, 50:61–76, 2006. Available from: http://www.ma.hw.ac.uk/~iain/research/GLAM/CSDA_2006/CSDA.html.
- [12] George Sterman et al. (CTEQ Collaboration). Handbook of perturbative qcd. *Rev. Mod. Phys.*, 67(1):157–248, Jan 1995. doi:10.1103/RevModPhys.67.157.
- [13] Felix Aharonian et al. (H.E.S.S. Collaboration). Primary particle acceleration above 100 TeV in the shell-type Supernova Remnant RX J1713.7-3946 with deep H.E.S.S. observations. *Astron. Astrophys.*, 464:235–243, 2007. arXiv:astro-ph/0611813v1, doi:10.1051/0004-6361:20066381.
- [14] Abraham Achterberg et al. (IceCube Collaboration). First year performance of the IceCube neutrino telescope. *Astroparticle Physics*, 26(3):155–173, 2006. doi:10.1016/j.astropartphys.2006.06.007.
- [15] Markus Ackermann et al. (IceCube Collaboration). Optical properties of deep glacial ice at the South Pole. *J. Geophys. Res.*, 111(D13):1–26, Jan 2006. doi:10.1029/2005JD006687.
- [16] Rasha Abbasi et al. (IceCube Collaboration). 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, 2009. doi:10.1016/j.nima.2009.01.001.

- [17] Claude Amsler et al. (Particle Data Group). Review of particle physics. *Physics Letters B*, 667(1), 2008. Available from: <http://pdg.lbl.gov/>.
- [18] Hans Gerd Evertz. The loop algorithm. *Advances in Physics*, 52:1, 2003. arXiv:cond-mat/9707221, doi:10.1080/0001873021000049195.
- [19] Tony A. Gabriel, Don E. Groom, P.K. Job, Nikolai V. Mokhov, and G.R. Stevenson. Energy dependence of hadronic activity. *Nuclear Instruments and Methods in Physics Research A*, 338:336–347, 1994. doi:10.1016/0168-9002(94)91317-X.
- [20] Raj Gandhi, Chris Quigg, Mary Hall Reno, and Ina Sarcevic. Ultrahigh-energy neutrino interactions. *Astropart. Phys.*, 5:81–110, 1996. arXiv:hep-ph/9512364, doi:10.1016/0927-6505(96)00008-4.
- [21] Raj Gandhi, Chris Quigg, Mary Hall Reno, and Ina Sarcevic. Neutrino interactions at ultrahigh energies. *Phys. Rev. D*, 58(9):093009, Sep 1998. doi:10.1103/PhysRevD.58.093009.
- [22] Andrew Gelman, Gareth O. Roberts, and Wally R. Gilks. Efficient metropolis jumping rules. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, editors, *Bayesian Statistics 5: Proceedings of the Fifth Valencia International Meeting, June 5-9, 1994*. Oxford Univ. Press, 1996. Available from: <http://www.stat.columbia.edu/~gelman/research/published/baystat5.pdf>.
- [23] W. Keith Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. Available from: <http://www.jstor.org/stable/2334940>.
- [24] John David Jackson. *Classical Electrodynamics*. John Wiley Sons, 3 edition, 1999.
- [25] Marek Kowalski. *Search for Neutrino-Induced Cascades with the AMANDA-II Detector*. PhD thesis, Humboldt-Universität zu Berlin, January 2004.
- [26] John G. Learned and Karl Mannheim. High-energy neutrino astrophysics. *Annual Review of Nuclear and Particle Science*, 50(1):679–749, 2000. doi:10.1146/annurev.nucl.50.1.679.
- [27] Johan Lundberg, Predrag Miočinić, Kurt Woschnagg, Thomas Burgess, Jenni Adams, Stefan Hundertmark, Paolo Desiati, and Peter Niessen. Light tracking through ice and water—scattering and absorption in heterogeneous media with photonics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 581(3):619 – 631, 2007. arXiv:arXiv:astro-ph/0702108v2, doi:10.1016/j.nima.2007.07.143.
- [28] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953. doi:10.1063/1.1699114.
- [29] Eike Middell. Reconstruction of cascade-like events in icecube. Diploma thesis, Humboldt-Universität zu Berlin, July 2008.
- [30] Dario Motta and Stefan Schönert. Optical properties of bialkali photocathodes. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 539(1-2):217 – 235, 2005. doi:10.1016/j.nima.2004.10.009.

- [31] Radford M. Neal. Probabilistic inference using markov chain monte carlo methods. Technical Report CRG-TR-93-1, Dept. of Computer Science, University of Toronto, 1993. Available from: <http://www.cs.toronto.edu/~radford/review.abstract.html>.
- [32] Dirk Pandel. Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal-Neutrinooteleskop NT-72. Diploma thesis, Humboldt-Universität zu Berlin, 1996.
- [33] Bernhard Voigt. *Sensitivity of the IceCube Detector for Ultra-High Energy Electron-Neutrino Events*. PhD thesis, Humboldt-Universität zu Berlin, April 2008. Available from: <http://edoc.hu-berlin.de/dissertationen/voigt-bernhard-2008-07-16/PDF/voigt.pdf>.
- [34] Chris Wendt. Waveform calibration. In *IceCube Fall Collaboration Meeting (Berlin)*, September 2009. Available from: <https://docushare.icecube.wisc.edu/dsweb/Get/Document-51318/WaveformCalibration-Berlin-Sep2009.pdf>.
- [35] Chris Wendt, Shigeru Yoshida, and Keiichi Mase. Calibration and characterization of the IceCube photomultiplier tube. 2009.
- [36] Nathan Whitehorn. Photonics binning issue update. In *IceCube Spring Colloboration Meeting 2009 (Madison, WI, USA)*, April 2009.
- [37] Christopher Wiebusch. *The Detection of Faint Light in Deep Underwater Neutrino Telescopes*. PhD thesis, RWTH Aachen, December 1995. Available from: <http://web.physik.rwth-aachen.de/~wiebusch/Publications/Various/phd.ps.gz>.
- [38] Christopher Wiebusch. Physics capabilities of the IceCube DeepCore detector. In *Proceedings of the 31st ICRC, Łódź, Poland*, July 2009. arXiv:0907.2263v1.
- [39] Kai Zuber. *Neutrino physics*. Institute of Physics Publishing, Philadelphia, 2003.

A Total event charge as a function of energy

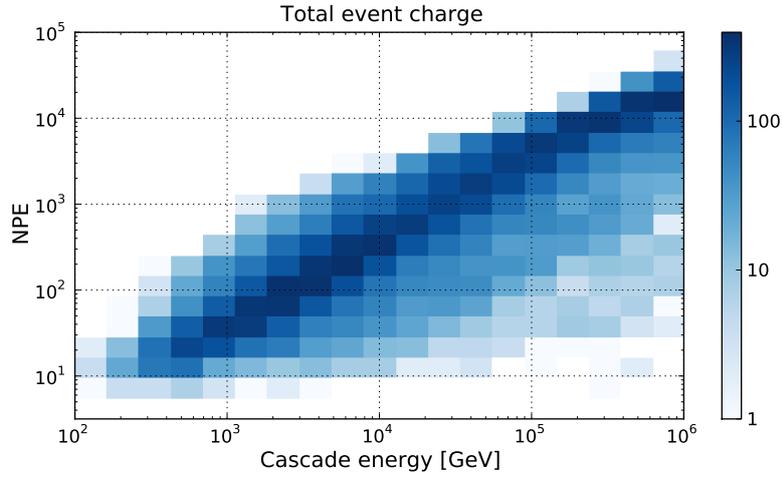


Figure 34: Total event charge as a function of cascade energy for the full simulation, which included a realistic simulation of hadronic cascades and an inhomogeneous ice model. The selection includes only fully-contained events with no hits on the outer strings of the detector or the first or last DOM on a string.

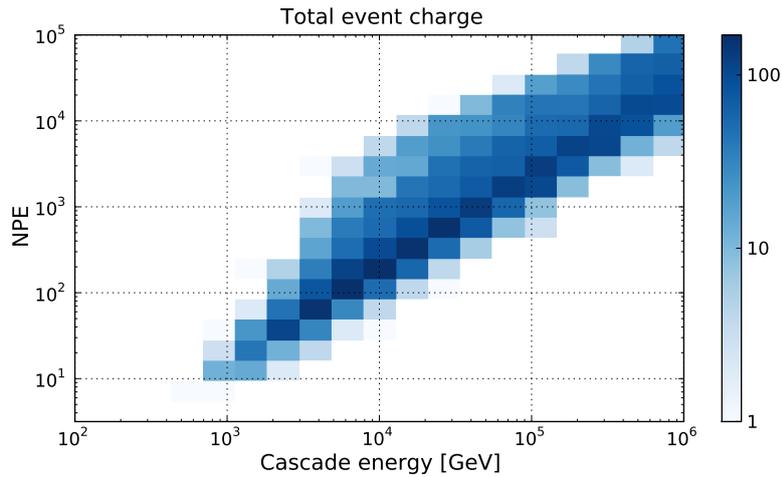


Figure 35: Total event charge as a function of cascade energy for the simplified simulation using point-like electromagnetic cascades. The events were subject to the same geometric cut as in Figure 34.

B Further resolution plots

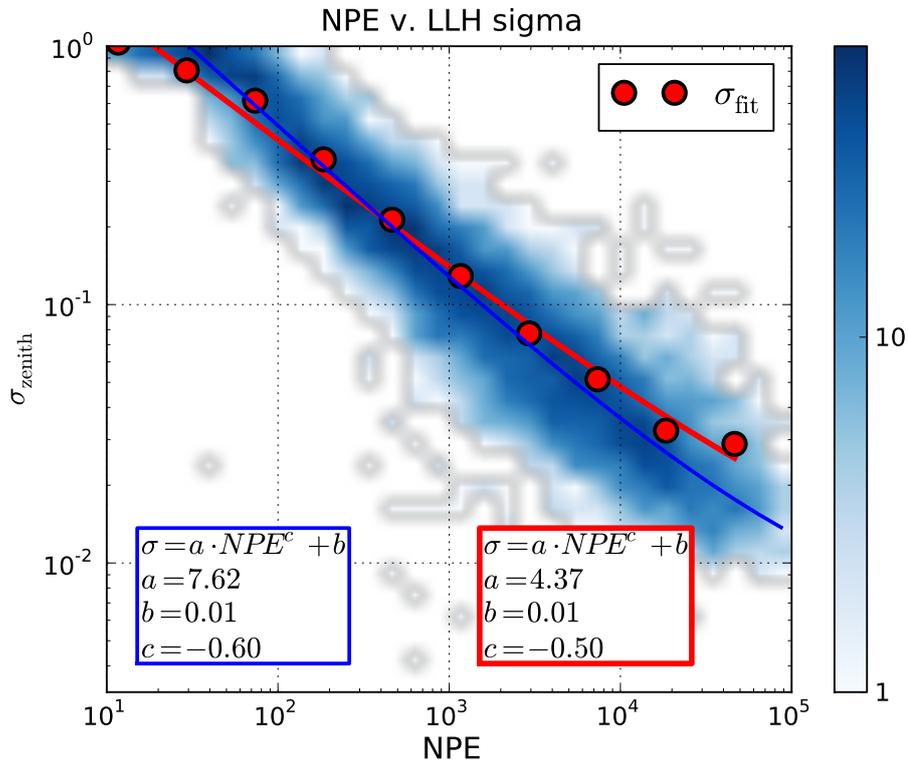


Figure 36: Resolution in zenith angle θ (σ_{zenith} , in radians) as a function of total event charge (NPE) for point-like cascades simulated with realistic electronics in bulk ice, with un-hit DOMs included in the likelihood calculation.

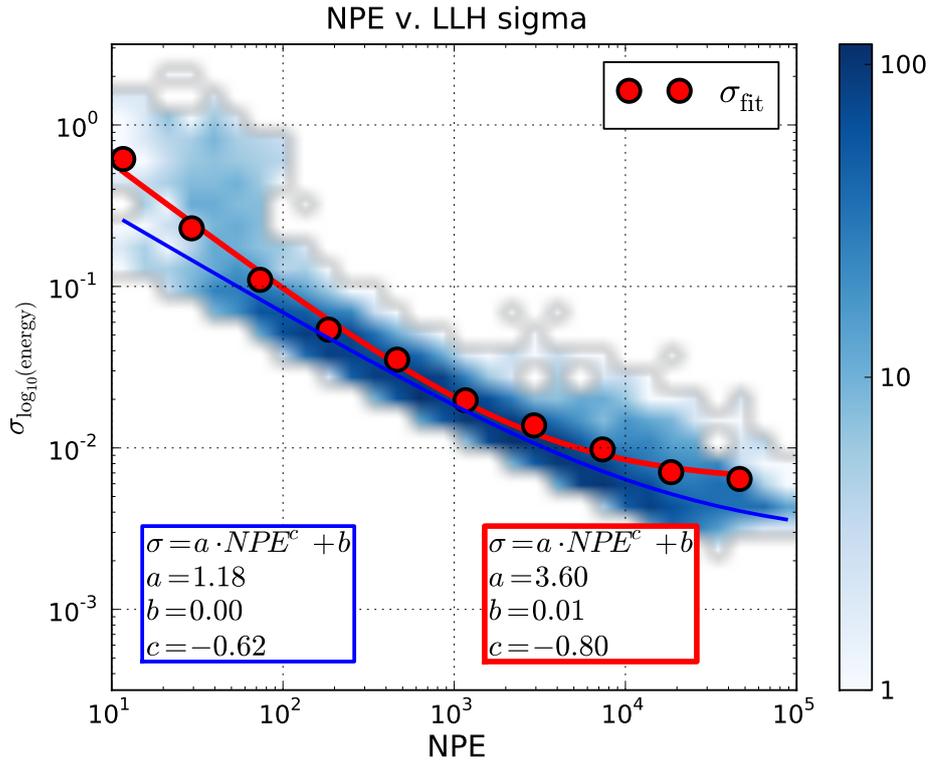


Figure 37: Resolution in the logarithm of the cascade energy E ($\sigma_{\log_{10} \text{energy}}$, in $\log \frac{E}{\text{GeV}}$) as a function of total event charge (NPE) for point-like cascades simulated with realistic electronics in bulk ice, with un-hit DOMs included in the likelihood calculation.

Zusammenfassung

In dieser Diplomarbeit handelt es sich um die Verbesserung der Rekonstruktion von elektromagnetischen Kaskaden im Neutrinoobservatorium ICECUBE mittels Markov-Ketten.

Bei Ereignissen dieser Klasse werden Schauer von hochenergetisch geladenen Teilchen erzeugt, die wegen ihrer relativistischen Geschwindigkeit Cherenkov-Licht im antarktischen Gletschereis ausstrahlen. Die Cherenkov-Photonen werden von einem Raster Photovervielfacherrohren (PMTs) im Eis in eine Reihe von Spannungspulsen umgewandelt die vor Ort digitalisiert werden. Aus diesen Daten werden die Ankunftszeiten der Photonen rekonstruiert, welche die Basis für die Rekonstruktion des Ereignisses im Eis bilden.

Mit Hilfe einer Beschreibung der Streuung und Absorption der Photonen im Eis, sowie der Reaktion der PMTs und der Ausleseelektronik auf einfallende Photonen kann ein statistisches Modell des Detektors definiert werden. Dieses statistische Modell beschreibt die Wahrscheinlichkeit, mit der eine bestimmte Hypothese (festgelegt durch die Position und die Ursprungszeit des Schauers sowie dessen Richtung und Energie) zu verschiedenen Datensätzen (Scharen von rekonstruierten Photonenankunftszeiten) führt. Statt als Funktion der Datenparameter mit fixierter Hypothese kann dies auch als Funktion der Hypothesenparameter mit fixierten Daten betrachtet werden; so betrachtet wird sie "Likelihood-Funktion" genannt. Die "Likelihood" einer Hypothese ist der Wert der Likelihood-Funktion angewandt auf die jeweilige Hypothese. Im Rahmen des statistischen Modells ergibt das Verhältnis zweier Likelihood-Werte ein Maß dafür, wie stark eine Hypothese von den Daten unterstützt wird im Verhältnis zur zweiten Hypothese. Die von den Daten am besten unterstützte Hypothese kann durch numerisches Maximieren der Likelihood-Funktion ermittelt werden. Darüber hinaus enthält die Form des Maximums der Likelihood-Funktion Informationen über die minimale Auflösung der Rekonstruktion.

In einer früheren Diplomarbeit [29] wurde eine Likelihood-Rekonstruktion für kaskadenartige Ereignisse in ICECUBE entwickelt. Meine Arbeit beschäftigt sich damit, das Verhalten der Vertex-, Richtungs-, und Energieauflösung dieser Rekonstruktion mittels Markov-Ketten zu untersuchen.

Eine Markov-Kette ist eine Reihe von gegenseitig abhängigen Stichproben ("Schritte") aus einem Parameterraum in dem der Einfluss der vorausgehenden Schritte auf die Verteilung der Stichproben im nächsten Schritt allein durch den letzten Schritt vermittelt werden. Es lassen sich Markov-Ketten erstellen in dem sich die Verteilung der Stichproben nach hinreichend vieler Schritte einer willkürlichen Zielverteilung nähert. In dieser Diplomarbeit wird eine Markov-Kette verwendet, um anhand eines Datensatzes eine Schar Stichproben aus dem 7-dimensionalen Parameterraum der Kaskadenhypothesen zu erzeugen, deren Verteilung sich der Likelihood-Funktion nähert. Im gauß'schen Grenzwert ergibt die Varianz dieser Verteilung in verschiedenen Dimensionen die minimale Auflösung der Rekonstruktion in der jeweiligen Dimension.

Diese Methode wurde anhand simulierter Daten getestet und die ermittelte Auflösung mit der Verteilung der Rekonstruktionsresidua verglichen. Nach ersten Versuchen konnten die beiden Auflösungen sich um mehr als eine Größenordnung unterscheiden. Dies wies darauf hin, dass das statische Modell der vorhandenen Likelihoodfunktion die Simulation nicht ausreichend beschrieben hat. Im Rest der Arbeit wurde die Simulation bis auf das Nötigste vereinfacht bis die Rekonstruktion selbstkonsistent wurde. Danach wurde sie wieder teilweise aufgebaut und gleichzeitig passende Korrekturen zur Likelihood-Funktion eingeführt. Außerdem wurde eine neue Parametrisierung der Cherenkovlicht-Verteilung angewendet, um aus linearer Interpolation stammende Fehler zu vermeiden. Mit diesen Verbesserungen konnten die Rekonstruktionsresidua verkleinert werden und auch zuverlässig anhand der Markov-Kette voraus gesagt werden.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Ort und Datum

Unterschrift

