Studies of the Reconstruction of Cascade-Like Events in PINGU

Masterarbeit in Physik

von

Thomas Ehrhardt

angefertigt im

Physikalischen Institut
Nussallee 12
53115 Bonn

vorgelegt der

Mathematisch-Naturwissenschaftlichen Fakultät

der Universität Bonn

Dezember 2014

1. Gutachter: Herr Prof. Dr. Marek Kowalski
2. Gutachter: Herr Prof. Dr. Jochen Dingfelder

Contents

1. Introduction

Now that a large number of experiments, measuring neutrinos from different sources such as reactors, accelerators, the sun or the Earth's atmosphere, have provided compelling evidence for the existence of non-vanishing neutrino masses via oscillations, one of the key questions that has yet to be answered is that of the mass ordering. Whether it is normal or inverted is not only of practical importance to experiments that address such fundamental uncertainties as the absolute masses of neutrinos or their Dirac or Majorana nature, but can also help constrain theoretical models at the GUT mass scale [1].

The Precision IceCube Next Generation Upgrade (PINGU), which is planned as an infill-array to the existing IceCube neutrino detector, located at the South Pole Station, has the potential to establish the neutrino mass hierarchy with high significance at relatively short time scale. In order to reach this goal, it will exploit characteristic modifications of the oscillation probabilities of atmospheric neutrinos as they traverse the Earth on their path towards the detector.

Only very recently, the IceCube collaboration reported on measurements of the atmospheric neutrino mixing parameters $|\Delta m^2_{32}|$ and $\sin^2\theta_{23}$, which are already “comparable in precision to those of dedicated oscillation experiments” [2], not least because of the denser instrumentation of IceCube's DeepCore region, where the Antarctic ice is clearest. With its even closer spacing of optical sensors, PINGU will be able to efficiently detect atmospheric neutrinos with energies above a few GeV and further constrain the values of the atmospheric oscillation parameters [1]. Reduced uncertainties on $|\Delta m^2_{32}|$ and $\sin^2\theta_{23}$, in turn, will directly benefit the neutrino mass hierarchy measurement.

Apart from a precise knowledge of systematics such as oscillation parameters, angular and energy resolutions are crucial. For PINGU, a dedicated universal reconstruction strategy has been devised that is adapted to the event topologies associated with low-energy neutrino interactions. After it had proven to yield good energy and direction estimates not only for track-like events, generally the result of charged current muon interactions, but also for cascades—which arise from charged current electron or tau neutrino and neutral current all-flavour neutrino interactions—the latter emerged as the main contribution to the overall sensitivity of PINGU [1]. Owing to their importance in the neutrino mass hierarchy determination, this work focusses on the reconstruction of cascade-like events in PINGU.

The most sophisticated reconstruction algorithms make use of a likelihood function that is maximised simultaneously in several dimensions, to find the set of parameters that best explains the observed data. The shape of the likelihood about its maximum then determines how tightly the best estimate is constrained, and thus yields the intrinsic resolution of the reconstruction on an event-by-event basis.

An attempt at extending an existing approach that yields semi-analytic resolution estimates for individual cascade-like events to the PINGU reconstruction strategy is introduced, and both are examined for their potential in predicting resolutions in scenarios of varying complexity. In the end, a simplified study is performed, which is based on the selection of well-reconstructed events to reduce the impact of systematics on the neutrino mass hierarchy significance in the cascade channel.

This work is structured as follows: chapter 2 briefly introduces the neutrino and measurements related to its mass. After the origin of the atmospheric neutrino flux is illustrated, the basics of neutrino oscillations in vacuum are laid out, and a qualitative picture of the modification of transition probabilities in (Earth) matter is given, which depends on the ordering of the neutrino masses. Afterwards, some key factors in the measurement of the mass hierarchy are discussed.

Chapter 3 first presents the physics processes that are relevant to neutrino detection in PINGU—especially those related to cascades—and then introduces the IceCube and PINGU detectors as well as the properties of the surrounding Antarctic ice.

In chapter 4, the most important steps in event simulation and low-level reconstruction are shortly explained, as they are drawn upon throughout this work.

Essentials of the method of maximum likelihood within the wider framework of parameter estimation are given in chapter 5, together with a prescription that allows one to obtain the (asymptotic) uncertainty of an estimate.

Chapter 6 details the constituting elements of the PINGU reconstruction strategy, and illustrates how reconstruction residuals can be estimated semi-analytically after a maximum likelihood reconstruction has been performed, based on tabulated light yields that have been fit with splines.

In chapter 7, the ideal relation between the likelihoods of the true and reconstructed event parameters is discussed, in order to subsequently be able to estimate the suitability of the likelihood formulation. Different simulation and reconstruction scenarios are created, with the aim to uncover inconsistencies which might affect the accuracy of a likelihood-based residual estimator.

An application of the estimator is shown in chapter 8, where it is utilised to perform quality cuts on a simulated data set of electron neutrino events in PINGU, so that the implications of different reconstruction resolutions for the neutrino mass hierarchy determination can be examined.

Chapter 9 will then conclude with a summary and a short outlook.

2. Neutrino Mass Hierarchy and Oscillations of Atmospheric Neutrinos

There are three neutrino flavours which interact weakly via charged (CC) or neutral currents (NC): electron, ν_e and $\bar{\nu}_e$, muon, ν_μ and $\bar{\nu}_\mu$, as well as tauon, ν_τ and $\bar{\nu}_\tau$. In CC weak interactions, each neutrino produces its charged lepton counterpart, i.e. ν_e ($\bar{\nu}_e$) the electron e^- (e^+), ν_μ ($\bar{\nu}_\mu$) the muon μ^- (μ^+), and ν_τ ($\bar{\nu}_\tau$) the tauon τ^- (τ^+).

In the famous Wu experiment conducted in 1956, C.S. Wu and coworkers [3] found that anti-neutrinos stemming from the β decay of polarised cobalt-60 nuclei were always emitted with positive helicity. As a result, in 1957 Yang and Lee [4] postulated that there were only two types of neutrino, the negative (left) helicity particle and the positive (right) helicity antiparticle. Since a massive spin one-half particle has two states of helicity, this two-component neutrino theory implied that the neutrino must be massless.

With the formulation of the maximally parity violating V-A theory [5], the weak force itself was soon found to act only on the left-handed states of quarks and leptons, which in principle allowed the neutrino to have a small mass. Nonetheless, since experiment did not indicate the opposite, left-handed massless neutrinos were incorporated in the minimal Standard Model [6].

There are two conservation laws that result from the assumption of vanishing neutrino masses. Firstly, individual lepton-flavour numbers, i.e electron-number, muon-number and tau-number, are conserved. Similarly, being the sum of individual flavour numbers, total lepton number is also conserved. As a consequence, in the framework of the minimal Standard Model, neutrino oscillations, i.e. in-flight transitions between different flavour neutrinos, and neutrinoless double beta decay $0\nu\beta\beta$ are forbidden [6].

In the late 1960's, first hints at the existence of neutrino oscillations, and thereby at nonzero neutrino masses, came from a chlorine radiochemical detector in the Homestake Mine in South Dakota [7], which used the neutrino capture reaction $\nu_e + {}^{37}\mathrm{Cl} \rightarrow {}^{37}\mathrm{Ar} + e^-$ to detect electron neutrinos stemming mainly from ${}^{8}\mathrm{B}$ decay taking place in the core of the sun. The fact that the observed neutrino flux was significantly below predictions made by the standard solar model (SSM) was the origin of the “solar-neutrino problem” [8], which today is interpreted as the transition of electron neutrinos into neutrinos of other flavours, caused by finite neutrino masses and neutrino mixing.

Several different experimental approaches have thus far lead to constraints on the absolute neutrino mass. The Troitsk [9] and Mainz [10] experiments, which measured the energy distribution of electrons in ${}^{3}H$ beta decay near the endpoint, found similar upper bounds on the mass of the electron anti-neutrino of $m_{\bar{\nu}_e} < 2.05\,\mathrm{eV}$ and $m_{\bar{\nu}_e} < 2.3\,\mathrm{eV}$ at 95 % CL, respectively. Measurements of the Cosmic Microwave Background (CMB) performed by the WMAP experiment, in combination with supernova data and data on galaxy clustering, result in upper limits on the sum of neutrino masses of $\sum_j m_j \lesssim 0.3$–$1.3\,\mathrm{eV}$, 95 % CL [11]. Most recently, the Planck Collaboration reported an upper limit of $\sum_j m_j \lesssim 0.23\,\mathrm{eV}$, 95 % CL [12], when adding baryoacoustic oscillation data to new measurements of the CMB temperature power spectrum and WMAP polarisation CMB data [8].

Beside the absolute mass scale, the ordering of the neutrino mass eigenstates ν_i, also called the *neutrino mass hierarchy*, is one of the remaining fundamental open questions, which can be resolved by studying atmospheric neutrinos. The following will therefore first provide a brief discussion of the production mechanisms of atmospheric neutrinos, before giving an overview over the formalism of neutrino oscillations in vacuum and modifications to the latter introduced in the presence of matter, especially during propagation through the Earth. Finally, the roles of some key parameters in the determination of the neutrino mass ordering with the envisaged PINGU detector are examined.

2.1. Atmospheric neutrinos

Interactions of primary cosmic rays within the Earth's atmosphere produce a large number of secondaries, whose decay leads to a steady flux of atmospheric neutrinos. The dominant decay processes are [13]

$$\begin{aligned} \pi^\pm &\rightarrow \mu^\pm \nu_\mu(\bar{\nu}_\mu) \\ K^\pm &\rightarrow \mu^\pm \nu_\mu(\bar{\nu}_\mu) \\ K_L &\rightarrow \pi^\pm e^\pm \nu_e(\bar{\nu}_e)\ , \end{aligned}$$

with possible further muon decays

$$\mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu \quad \text{and} \quad \mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu\ .$$

These decays can only occur, however, if the muon energy is below several GeV. Due to the characteristic height of around 15 km at which muons are produced, their decay length becomes too large otherwise.

Atmospheric electron and muon neutrinos stemming from the decay of charged pions and kaons constitute the *conventional* atmospheric neutrino flux, which is the dominant component up to energies of some 100 TeV [14]. It falls off as a power-law in energy, with a spectral index of $\gamma \approx 3.7$, as compared to $\gamma \approx 2.7$ for the primary cosmic ray flux. At higher energies, a different type of flux is expected to set in: the cosmic ray energy spectrum is "converted" into a *prompt* neutrino flux of equal spectral index via the decay of short-lived charmed mesons. This is the only mechanism that is perceived to produce a flux of atmospheric tau neutrinos [13].

Fig. 2.1 shows a compilation of predictions and measurements of the atmospheric ν_e and ν_μ fluxes, ranging from sub-GeV to PeV neutrino energies. A recent IceCube analysis [15] (red circles) probed the ν_μ flux at energies for which the spectrum is expected to become harder as the result of the prompt component, but no conclusive evidence was found due to relatively large uncertainties.

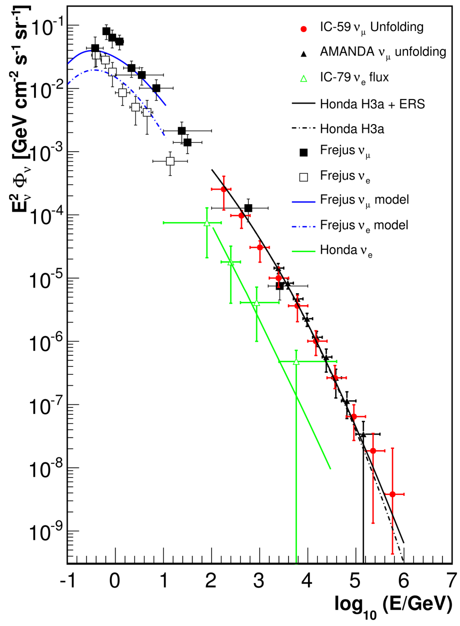

Figure 2.1.: Theoretical models and measurements of the atmospheric ν_μ and ν_e energy spectrum from the sub-GeV to the PeV range. See text for details. Figure taken from Ref. [15].

2.2. Vacuum neutrino oscillations

The mechanics of neutrino oscillations differ somewhat from those of flavour oscillations in the quark sector. On the one hand, neutrinos are stable and occur as free particles, and on the other they are always produced and detected as pure weak states through interactions of the weak force, while quarks start and end as states of defined mass [6].

The neutrino mixing matrix U is a unitary matrix that describes the relationship between the flavour eigenstates $|\nu_\alpha\rangle$ with $\alpha = e, \mu, \tau$ and the mass eigenstates $|\nu_i\rangle$:

$$|\nu_\alpha\rangle = \sum_{i=1}^{n} U_{\alpha i} |\nu_i\rangle \ . \tag{2.1}$$

The total number of massive neutrinos $n \geq 3$, depending on whether sterile neutrinos[1] exist, which mix with the flavour neutrinos [8]. Assuming $n = 3$, there exist three different mass splittings (only two of which are independent), Δm^2_{21}, Δm^2_{31}, and Δm^2_{32}, where $\Delta m^2_{ij} = m^2_i - m^2_j$, and three independent mixing parameters or Euler angles θ_{12}, θ_{13}, θ_{23}, which are used together with a CP violating phase δ_{CP} to parametrise the matrix U [16],

---

[1] These are hypothetical predominantly RH neutrinos or LH anti-neutrinos which do not couple to the weak $W^\pm$ and Z^0 gauge bosons [8].

$$\begin{aligned} \boldsymbol{U} &= \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} = \boldsymbol{U}^{(23)}\boldsymbol{U}^{(13)}\boldsymbol{U}^{(12)} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} , \end{aligned}$$

where $c_{ij} \equiv \cos\theta_{ij}$, $s_{ij} \equiv \sin\theta_{ij}$. The amplitude $A_{\alpha\to\beta}$ of finding a neutrino of flavour α produced with momentum $|\vec{p}| = k$ in a weak interaction process as one of a different flavour β at time t after production depends on the phases which are accumulated by the different mass eigenstates (that compose the original neutrino) during propagation [17],

$$A_{\alpha\to\beta} = \sum_i U_{\alpha i}U^*_{\beta i}e^{-im_i^2 t/(2k)} \ . \tag{2.2}$$

With $\delta_{CP} = 0$, i.e. CP conservation in the leptonic sector, the 3×3 matrix U becomes real, $\boldsymbol{U^*} = \boldsymbol{U}$, and from Eq. (2.2) follows for the transition probability

$$P(t)_{\alpha\to\beta} = |A_{\alpha\to\beta}|^2 = \sum_i U^2_{\alpha i}U^2_{\beta i} + 2\sum_{i,j;j>i} U_{\alpha i}U_{\alpha j}U_{\beta i}U_{\beta j}\cos(\Delta m^2_{ji}t/(2k)) \ . \tag{2.3}$$

Often neutrino oscillation experiments can be analysed in a simplified *2-neutrino mixing scenario*, in which Eq. (2.1) reduces to

$$|\nu_l\rangle = \cos\theta|\nu_1\rangle + \sin\theta|\nu_2\rangle \ , \quad |\nu_x\rangle = -\sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle \ . \tag{2.4}$$

Here, θ is the neutrino mixing angle in vacuum and ν_x represents a neutrino of another flavour, e.g. $l = \mu$ and $x = \tau$ [8]. Using the substitution L/E for t/k in Eq. (2.3), it is then straightforward to show that the survival and transition probabilities become

$$P(L)_{l\to l} = 1 - \sin^2 2\theta \sin^2\left(\frac{\Delta m^2}{4E}L\right) \ , \ \Delta m^2 = m_2^2 - m_1^2 > 0 \ , \text{ and} \tag{2.5}$$

$$P(L)_{l\to x} = 1 - P(L)_{l\to l} \ , \tag{2.6}$$

respectively. Thus, $P(\nu_l \to \nu_x)$ exhibits oscillatory dependence on the neutrino propagation distance L and energy E, while the factor $\sin^2 2\theta$ corresponds to the oscillation amplitude.

Suitable observation distances or *baselines* are best evaluated using the oscillation length $L_{\mathrm{osc}} = 4\pi E/(\Delta m^2)$, which is the propagation distance corresponding to one full oscillation.[2]

At one extreme, when the baseline $L \ll L_{\mathrm{osc}}$, even if the mixing angle θ is large, the oscillation will be at too early a stage to be observed, and $P(L)_{l\to l} \cong 1$, while $P(L)_{l\to x} \cong 0$. At the other extreme, $L \gg L_{\mathrm{osc}}$, it will only be possible to determine the average oscillation

---

[2] Given E in MeV and Δm^2 in eV^2, $L_{\mathrm{osc}} \cong 2.48\,\mathrm{m}\frac{E}{\Delta m^2}$.

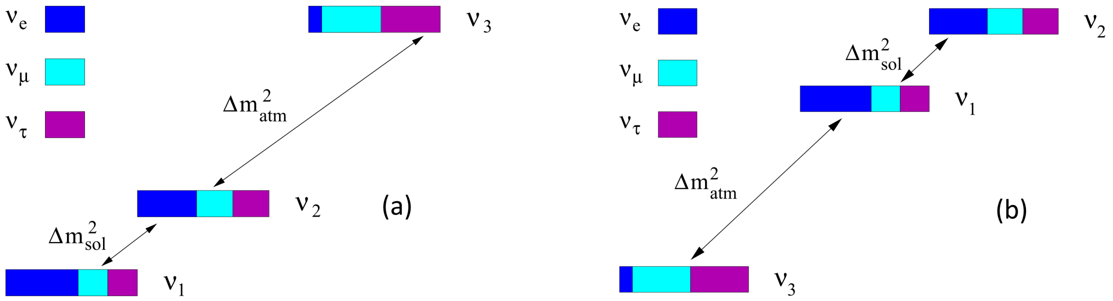

Figure 2.2.: Neutrino mass ordering under normal (a) and inverted (b) hierarchy and flavour admixtures of the mass eigenstates. See text for details. Figure taken from Ref. [18].

probabilities $\langle P_{l\to l}\rangle = 1 - 1/2\sin^2 2\theta$ and $\langle P_{l\to x}\rangle = 1/2\sin^2 2\theta$. In order for the $\nu_l \to \nu_x$ transition probability to become reasonably large, $P_{l\to x} \cong 1$, one requires $\sin^2 2\theta \cong 1$, and in addition that $L_{\text{osc}} \simeq 2\pi L$. For a given neutrino oscillation experiment, which detects neutrinos with (average) energy E at a distance L from their origin, the minimum resolvable mass squared difference is then determined by the condition $\Delta m^2_{\text{min}} \sim 2E/L$ [8].

2.3. Matter effect dependence on neutrino mass hierarchy

Within the picture of *3-neutrino mixing*, the two independent mass-squared differences are by convention chosen to be $\Delta m^2_{21} > 0$ and $\Delta m^2_{31} \neq 0$, with the absolute value $|\Delta m^2_{21}|$ of the former taken as the smaller of the two differences. While the absolute values obtained from different oscillation data imply $|\Delta m^2_{21}| \cong 7.5 \times 10^{-5}\,\text{eV}^2$ and $|\Delta m^2_{31}| \cong 2.5 \times 10^{-3}\,\text{eV}^2$ [8], it still remains unknown whether ν_3 is the most or least massive of the three neutrino mass eigenstates. The ordering (hierarchy) of the neutrino mass spectrum is referred to as *normal* in the former scenario (NH), $m_1 < m_2 < m_3$, and *inverted* in the latter (IH), $m_3 < m_1 < m_2$. These two possible mass hierarchies are depicted in Fig. 2.2, together with the flavour composition of the mass eigenstates.

Owing to the smallness of the element $|U_{e3}| = \sin\theta_{13}$, one can associate the mass-squared differences Δm^2_{21} and Δm^2_{31} (and the corresponding mixing angles) with the solar ν_e and dominant atmospheric ν_μ (and $\bar{\nu}_\mu$) oscillations, respectively. For this reason, Δm^2_{21} and Δm^2_{31} are also called solar and atmospheric neutrino mass-squared differences and denoted as $\Delta m^2_{21} \equiv \Delta m^2_{\text{sol}}$, $\Delta m^2_{31} \equiv \Delta m^2_{\text{atm}}$.

In matter, a significant enhancement of the oscillation probability can occur if there is a periodic change of density along the trajectory of the neutrino, the *parametric resonance*. Probabilites can exhibit further enhancement via the *MSW effect*, depending on the ordering of neutrino masses. A short discussion of these two mechanisms is given in the following.

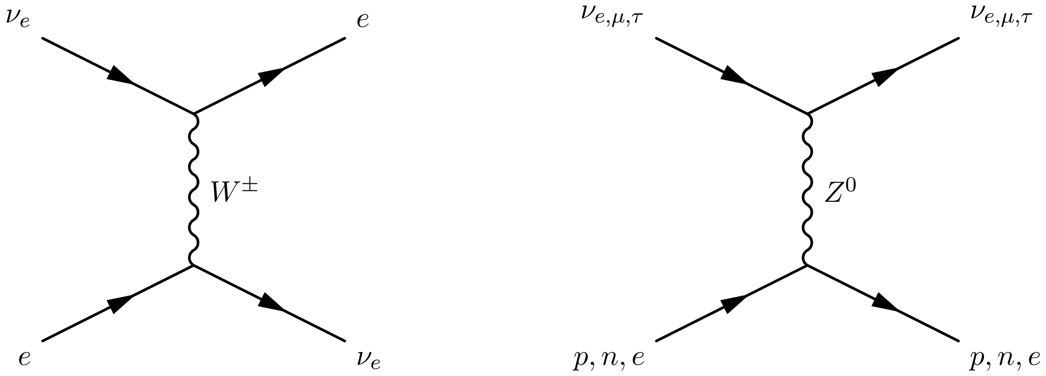

Figure 2.3.: Feynman diagrams for coherent elastic neutrino forward scattering on matter constituents. Only electron neutrinos exhibit charged current interactions (left), while neutral currents can mediate the interactions of neutrinos of all three flavours (right). See text for details.

Parametric Enhancement Parametric enhancement of oscillation probabilites can occur when atmospheric neutrinos traverse the Earth's core. In the *Preliminary Reference Earth Model* (PREM) [19], a number of spherical shells with different radii constitute the Earth's density profile. While the change in density as a function of radius within each shell proceeds relatively smoothly, the boundaries are characterised by rather sharp transitions. Since the density jumps between the Earth core and mantle are much larger than those between other shells, the density profile which is experienced by neutrinos crossing the core can be represented by only three layers, i.e. mantle-core-mantle [20].

The parametric resonance condition constrains the oscillation phases accumulated in the mantle and in the core, as defined e.g. in Ref. [20, Eq. (19)]. For one, these phases depend on the path lengths in the two layers of different (approximately constant) densities, and therefore on the zenith angle of the neutrino. But they are also crucially influenced by the values of the mixing parameters Δm^2_{32} and $\sin^2 2\theta_{13}$.

MSW Effect One cause for neutrino oscillations being modified in the presence of matter can be found in the two Feynman diagrams in Fig. 2.3, which show coherent forward elastic scattering processes of neutrinos on matter constituents. $W^\pm$ mediated CC interactions only occur for incident electron neutrinos, whereas neutrinos of any of the three flavours can scatter on protons, neutrons and electrons under the exchange of a Z^0 boson in neutral current interactions. The potential induced by these are identical in leading order, and thus do not alter survival and transition amplitudes. As a result of the CC interaction of electron neutrinos with the electrons, there is an effective potential $V = V_e^{\mathrm{CC}} = \sqrt{2} G_F N_e$, G_F and N_e being the Fermi constant and the electron number density in matter, respectively [18].

In a two-flavour scenario of $\nu_e \leftrightarrow \nu_\mu$ or $\nu_e \leftrightarrow \nu_\tau$ oscillations in matter with constant N_e, the dependence of $\sin^2 2\theta_m$ on N_e, where θ_m is the mixing angle in matter, will exhibit a resonance character. The MSW resonance condition is met when

$$N_e = N_e^{\mathrm{res}} = \frac{\Delta m^2 \cos 2\theta}{2E\sqrt{2}G_F} \;, \tag{2.7}$$

θ being the vacuum mixing angle. Alternatively, for given density N_e, the resonance occurs at the resonance energy E_{res}. Eq. (2.7) implies that, if $\Delta m^2 > 0$ (NH), the mixing in matter becomes maximal for $N_e = N_e^{\text{res}}$, even if the vacuum mixing angle θ is small. For oscillations of anti-neutrinos $\bar{\nu}_e \leftrightarrow \bar{\nu}_{\mu(\tau)}$, on the other hand, due to the replacement of N_e with $(-N_e)$, one finds the possibility of resonance enhancement exists only for $\Delta m^2 < 0$ (IH) [8].

Combined Picture Qualitatively, the trajectory of the neutrino as it passes through the Earth determines which of the two effects contributes to the modification of oscillation probabilities. For a zenith angle θ_ν such that only the mantle is traversed, $\cos\theta_\nu > -0.84$, oscillation probabilities can only be modified through the MSW resonance in the mantle, while for $\cos\theta_\nu < -0.84$ parametric and MSW resonance effects interfere. MSW resonances in the mantle and core will appear at around 7 GeV and 3 GeV, respectively, with the value at which parametric enhancement occurs lying in between [21].

Fig. 2.4 shows the oscillation probabilities $P(\nu_x(\bar{\nu}_x) \to \nu_\mu(\bar{\nu}_\mu))(x = e, \mu)$ extracted from Ref. [22], for both possible mass ordering schemes and two different neutrino zenith angles, as a function of neutrino energy. In the case of the vertically upgoing trajectory ($\cos\theta_\nu = -1$) and the NH (IH), there are modifications of the $\nu_e \to \nu_\mu$ ($\bar{\nu}_e \to \bar{\nu}_\mu$) transition probabilities for $E_\nu \lesssim 7\,\text{GeV}$ through parametric enhancement. Toward the highest energies shown, $E_\nu \gtrsim 15\,\text{GeV}$, the probability for ν_e to oscillate into ν_μ becomes increasingly small, and the different mass hierarchies become less and less distinguishable.

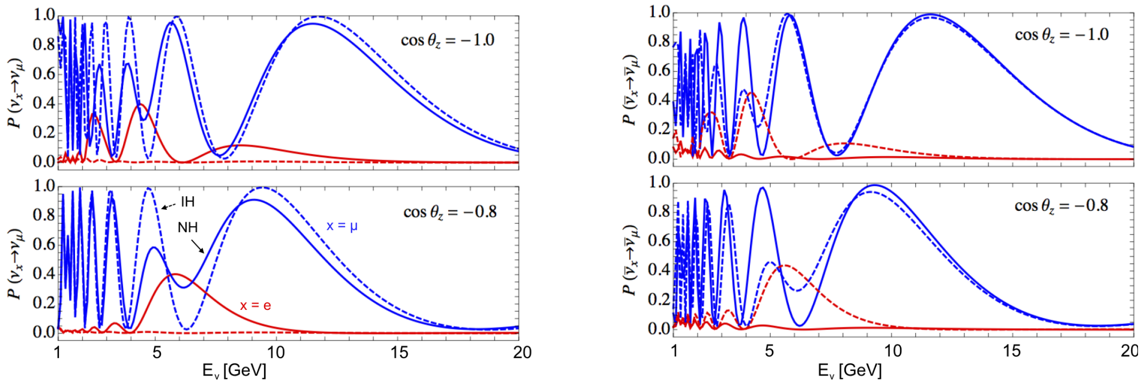

Figure 2.4.: Transition probabilities $\nu_e \to \nu_\mu$ (red graphs) and $\nu_\mu \to \nu_\mu$ (blue graphs) as a function of neutrino energy, for two different trajectories through the Earth. The solid lines correspond to the NH, while the dashed lines represent the IH. Left plots show neutrino, right plots show anti-neutrino oscillation probabilities. Figure adapted from Ref. [22].

2.4. Measurement of the neutrino mass hierarchy asymmetry

A non-magnetised neutrino detector such as PINGU has no charge discrimination power, i.e. it will only be able to measure the sum of neutrino and anti-neutrino event rates. Fig. 2.4 shows that, to a first approximation, the transition probabilities for neutrinos under the NH are the same as those for anti-neutrinos under the IH.[3] For a given flavour, a net asymmetry between NH and IH remains nevertheless, because on the one hand the CC cross sections for neutrinos in the relevant energy range of a few GeV exceed the corresponding anti-neutrino cross sections by a factor of about two (see Fig. 3.1), and on the other hand the flux of atmospheric neutrinos is larger than that of anti-neutrinos by about 30 % [21].

A number of parameters that might impact the measurement of the asymmetry and reduce the significance with which the wrong mass hierarchy can be rejected are related to the detector itself. In this regard, the neutrino energy and zenith angle estimation performance play a key role. As is demonstrated quantitatively in Ref. [22], accurate reconstructions are necessary to obtain a high sensitivity to the NMH, a major fraction of which arises from rather clearly confined regions in the $(E_\nu, \cos\theta_\nu)$ plane; this aspect will also be examined in the course of this work. Further, detector related systematics include uncertainties in energy calibration and effective areas. Theoretical uncertainties in (anti-)neutrino cross sections are partially degenerate with uncertainties in the spectral index and relative normalisation of the atmospheric neutrino fluxes (see Refs. [1, 23]).

Apart from detector-related systematics, it is also necessary to study the impacts of the different oscillation parameters, especially of Δm^2_{31} and θ_{23}. Uncertainties in these parameters can significantly reduce the significance of the hierarchy determination by giving rise to patterns in the $(E_\nu, \cos\theta_\nu)$ plane that resemble actual neutrino mass hierarchy effects [22].

---

[3] They would be identical if it was not for non-zero 1-2 mass splittings and mixing [22].

3. Neutrino Detection in Ice

3.1. Interaction cross sections of neutrinos at GeV energies

Neutrino interaction with nuclei in a target medium can proceed either via *charged currents* (CC) or via *neutral currents* (NC). While the former process is characterised by the exchange of a charged W-boson, the latter is mediated by the neutral Z-boson. With increasing energy, the neutrino starts resolving the individual constituents of the target.

At intermediate energies $E_\nu \sim 0.1$–$20\,\mathrm{GeV}$, neutrino scattering comprises the three main mechanisms (quasi)elastic scattering, resonance production and deep inelastic scattering [24]. In (quasi)elastic interactions, neutrinos scatter off an entire nucleon, thereby separating one or more nucleons from the target. Resonance production refers to a struck nucleon being converted into an excited state, resulting in numerous possible final states. Deep inelastic scattering is by far the dominant process at neutrino energies exceeding $20\,\mathrm{GeV}$. Here, the neutrino interacts with a quark in the nucleon. Depending on the type of boson exchanged, the particles in the final state differ [13]:

$$\nu_l N \rightarrow l^- X \quad \text{and} \quad \bar{\nu}_l N \rightarrow l^+ X \quad (\mathrm{CC})\;, \tag{3.1}$$

$$\nu_l N \rightarrow \nu_l X \quad \text{and} \quad \bar{\nu}_l N \rightarrow \bar{\nu}_l X \quad (\mathrm{NC})\;, \tag{3.2}$$

where $l = e, \mu, \tau$ denotes the lepton flavour, N is either a proton, neutron or an isoscalar target, and X is the hadronic final state. Both types of interactions comprise a system of final state hadrons, which give rise to a hadronic particle shower, described in Sec. 3.4.2. The lepton l, on the other hand, only emerges in CC interactions of a neutrino with corresponding flavour.

Fig. 3.1 shows the total CC cross sections for both neutrinos and anti-neutrinos as a function of energy, together with the contributions from the relevant individual processes alluded to above. While anti-neutrino cross sections are typically only half as large as their corresponding neutrino counterparts, the linear dependence on energy approached above some ten GeV is expected for scattering on point-like quarks [8].

3.2. Ionisation, bremsstrahlung and pair production

Depending on the type and energy of a particle passing through matter, different interaction processes, manifesting in a loss of energy, may play a role. This section will therefore give a brief overview over those most relevant to the detection of secondaries produced in (CC) interactions of neutrinos with GeV energies, based on Ref. [25].

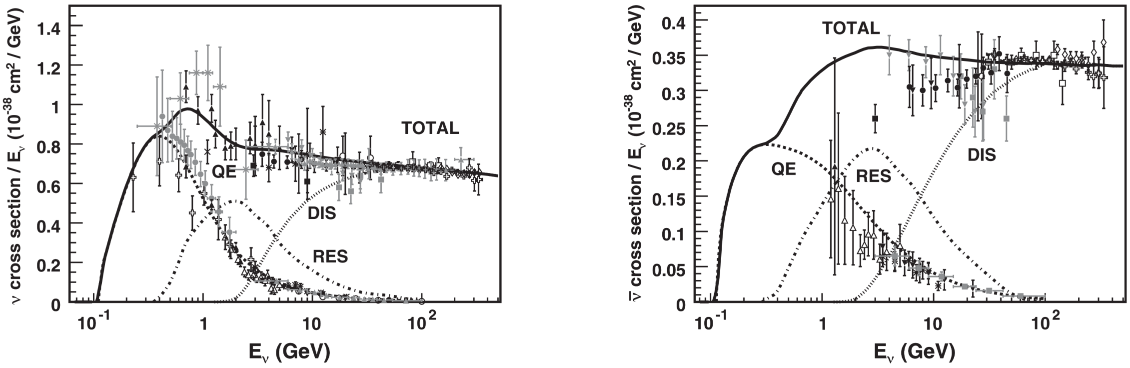

Figure 3.1.: Total neutrino (left) and anti-neutrino (right) CC cross sections per nucleon divided by neutrino energy, and individual contributions from resonance production (RES), quasielastic (QE) and deep inelastic scattering (DIS), at intermediate and high energies. See text for details. Figures taken from Ref. [24].

3.2.1. Ionisation by heavy particles

Ionisation or atomic excitation energy losses of fast charged particles with speed $v = \beta c$ are usually small ($< 100\,\mathrm{eV}$ in $90\,\%$ of cases). Here, heavy particles need to be considered separately from positrons and electrons due to kinematics, spin, and the identity of the latter with the atomic electrons. At intermediate energies $0.1 < \beta\gamma < 1000$ and charge numbers Z, the mean energy loss rate can be well-approximated by the Bethe formula,

$$\left\langle -\frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2 m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] . \tag{3.3}$$

Here, $K = 4\pi N_A r_e^2 m_e c^2$ is a constant, z the charge of the projectile in units of the elementary charge e, $T_{\max}$ the maximum kinetic energy that can be transferred to a free electron in a single collision, I the mean excitation energy in eV, and $\delta(\beta\gamma)$ a density effect correction. In a given material, $\langle -dE/dx \rangle$ can be viewed as a function of β alone. It typically exhibits a broad minimum, whose position shifts depending on Z.

For muons in ice, the mean energy loss rate at minimum ionisation is approximately $1.8\,\mathrm{MeV\,cm^{-1}}$, and in the energy range of interest, radiative processes, i.e. bremsstrahlung, pair production and photonuclear interactions, only contribute a negligible fraction to energy loss [26].

3.2.2. Photon and electron energy loss

For electrons and positrons at very low energies, ionisation is the primary energy loss mechanism. As it turns out, at similar values of γ, the resulting mean energy loss rates do not deviate strongly from those of heavier particles. Above some ten MeV in most materials, losses due to bremsstrahlung dominate. Bremsstrahlung is emitted by a moving charged particle in the Coulomb-field of a nucleus or an atomic electron.

Two important quantities describing the ionisation and bremsstrahlung energy losses of electrons and positrons are *radiation length* and *critical energy*. Both are also of importance in characterising the longitudinal development of electromagnetic cascades (cf. Sec. 3.4.1 below).

Radiation length The radiation length X_0 is a characteristic measure of the amount of matter traversed by both high-energy electrons emitting bremsstrahlung and high-energy photons producing electron-positron pairs. While in the former case it corresponds to the mean distance an electron needs to travel before its energy has decreased to $1/e$ of its initial value, in the latter it is equivalent to 7/9 of the mean free path before a photon creates an e^+e^- pair in the vicinity of a nucleus. In both cases, the energy loss rate scales approximately linearly with energy, and thus

$$\left\langle -\frac{dE}{dX} \right\rangle_\mathrm{i} \simeq c_\mathrm{i} \frac{E}{X_0} \ , \tag{3.4}$$

where $i =$ brems, pair and $c_\mathrm{brems} = 1$, $c_\mathrm{pair} = 7/9$. X_0 can be parametrised as a function of the medium's charge number Z; see Ref. [25, Eq. (31.26)]. Ref. [26] states a radiation length of $X_{0,\mathrm{ice}} = 36.08\,\mathrm{g\,cm^{-2}}$ for ice, which translates into a mean free distance of $39.65\,\mathrm{cm}$ assuming a density of $\rho_\mathrm{ice} = 0.91\,\mathrm{g\,cm^{-3}}$.

Critical energy Using the approximation of the energy loss due to bremsstrahlung given in Eq. (3.4), the critical energy E_c corresponds to the electron energy at which the loss rates due to ionisation and bremsstrahlung are equal, but also to the energy at which the electron energy and the ionisation energy loss per radiation length X_0 are equal. A simple fit to the critical energy in solids exists, which does not take into account existing dependencies on A, I, and other factors:

$$\text{solids:} \qquad E_c \simeq \frac{610\,\mathrm{MeV}}{Z + 1.24} \ . \tag{3.5}$$

The critical energies in ice obtained from Ref. [26] are $E^{e^-}_{c,\mathrm{ice}} = 78.60\,\mathrm{MeV}$ and $E^{e^+}_{c,\mathrm{ice}} = 76.51\,\mathrm{MeV}$.

3.3. Cherenkov radiation

A charged particle of mass m radiates Cherenkov light if the velocity $v = \beta c$ with which it propagates through a dielectric medium with refractive index n exceeds the local phase velocity of light c/n; the kinetic energy threshold is thus defined by

$$E_t = mc^2 \left(\frac{1}{\sqrt{1 - n^{-2}}} - 1 \right) \ . \tag{3.6}$$

Photons are emitted into a cone whose vertex is located at the position of the moving charged particle; its opening half-angle is called the *Cherenkov angle* and is given by

$$\theta_C = \arccos\left(\frac{1}{\beta n}\right) \ . \tag{3.7}$$

In dispersive media $n = n(\omega)$, so the relations given above hold for radiation of frequency ω under the condition that $v > c/n(\omega)$.

For a relativistic ($\beta = 1$) source in ice, whose refractive index is $n_{\text{ice}} \approx 1.33$ [27], Eq. (3.7) evaluates to $\theta_C \approx 41°$. The kinetic energy threshold for the emission of Cherenkov radiation ($\theta_C = 0$) for an electron in ice is $E_t \approx 0.26\,\text{MeV}$, while for a muon and a proton the thresholds are 54 and 475 MeV, respectively.

The Frank-Tamm formula quantifies the number of Cherenkov photons per unit path length and unit wavelength interval radiated by a particle with charge ze [25]:

$$\frac{\mathrm{d}^2 N}{\mathrm{d}x\,\mathrm{d}\lambda} = \frac{2\pi\alpha z^2}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right) \ , \tag{3.8}$$

where $\alpha \approx 1/137$ is the fine structure constant. In the 300–500 nm wavelength band, around 250 optical photons per cm of path length are produced by a particle of unit charge travelling at $\beta = 1$ through ice [27]. Large photomultiplier sensitivities (see Sec. 3.5.2) and a small intrinsic absorptivity of pure ice in this wavelength range [28] benefit the detection of charged particles via the considerable number of optical Cherenkov photons they emit.

While many high-energy cosmic ray and neutrino detectors make use of the Cherenkov effect, it only contributes a small amount to the energy loss of charged particles, i.e. around 0.5 % of that due to ionisation for a minimum-ionising particle [29].

3.4. Modelling of particle showers

When considering the physics of showers—cascades of secondary particles initiated by the interaction of a high-energy primary with matter constituents—one needs to distinguish between those produced by particles interacting electromagnetically and those originating from hadrons, because of the different forces and processes involved in each case.

3.4.1. Electromagnetic cascades

According to the concept of critical energy introduced in Sec. 3.2.2, an electron with energy $E > E_c$ subject to an electric field, as generated for example by the atoms in a target medium, will emit bremsstrahlung photons, which in turn can produce electron-positron pairs, which radiate further photons and so on; an electromagnetic cascade is initiated.

A description of the basic features of the longitudinal development of an electromagnetic shower is given within the Heitler model [30] (originally Carlson & Oppenheimer [31]): there is only one type of particle with energy E and fixed interaction length X_0; at each interaction two new particles with energy $E/2$ are produced (Figure 3.2). The number of particles N present in the shower at a given depth $z = n \cdot X_0$, where n denotes the number

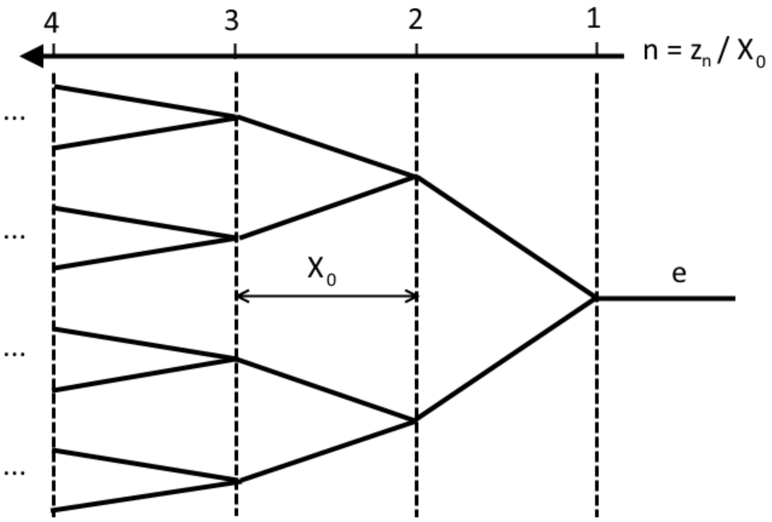

Figure 3.2.: Simple Heitler model describing the longitudinal development of an electromagnetic cascade up to its maximum. See text for details.

of generations or consecutive interactions, reads $N(z) = 2^n = 2^{z/X_0}$. Assuming the energy of the primary (shower-inducing) particle is E_0, each particle belonging to generation n has an energy $E(z) = E_0/2^n$.

The depth at which the number of particles in the cascade is at a maximum N_max follows from the condition $E = E_c$, since as soon as electron energies fall below the critical energy, their energy is dissipated by ionisation and excitation. As a consequence, particle-multiplication comes to a halt and the shower starts to subside. Hence,

$$N_\mathrm{max} = \frac{E_0}{E_c} \qquad \text{and} \qquad z_\mathrm{max}(E_0) \propto X_0 \ln\left(\frac{E_0}{E_c}\right) \ . \tag{3.9}$$

Realistic Monte Carlo simulations performed in Ref. [27], employing full tracking of the shower constituents, confirm the scaling of the shower maximum with the logarithm of the primary energy as predicted by this simple model.

Charged cascade constituents with energies above the Cherenkov threshold (3.6) will emit a number of Cherenkov photons that is proportional to their track length, which follows from Eq. (3.8). One should also note that the photon yield depends on the velocity β, and is reduced by the factor

$$\frac{n^2 - \beta^{-2}}{n^2 - 1} \tag{3.10}$$

with respect to that of a $\beta = 1$ charge. Eq. (3.10) then determines the total *effective track length* $\hat{l}$ of a cascade, whose dependence on the primary energy E_0 has been found to be linear in different studies [27, 32, 33], with the most recent one reporting a constant of proportionality of $\sim 532\,\mathrm{cm\,GeV^{-1}}$ and extremely small shower-to-shower fluctuations at a less than 1 % level already for energies in the GeV range [27].

A reasonably accurate description of the mean longitudinal energy deposition profile of an electromagnetic cascade can be given using a gamma distribution [25]:

$$\frac{\mathrm{d}E}{\mathrm{d}n} = E_0\, b\, \frac{(bn)^{a-1} e^{-bn}}{\Gamma(a)} \ , \tag{3.11}$$

where a and b are characteristic dimensionless constants. Ref. [27] finds a logarithmic increase $a = \alpha + \beta \log_{10}(E_0/1\,\mathrm{GeV})$, with $\alpha \approx 2.02$, $\beta \approx 1.45$ for an initial e^-, while $b \approx 0.63$ shows no energy dependence. The depth of maximum energy deposition follows as

$$n_{\max} = \frac{a-1}{b} \ . \tag{3.12}$$

According to this parametrisation, using the mean free distance in ice given in Sec. 3.2.2, the average maximum energy deposition of an electromagnetic cascade brought about by a primary electron of energy $E_0 = 10\,\mathrm{GeV}$ will occur at a depth of $z_{\max} \approx 1.55\,\mathrm{m}$.

Finally, the angular distribution of Cherenkov photons emitted during the development of the cascade is found to be quite broad, albeit with a pronounced maximum at the Cherenkov angle, and is independent of primary energy (Fig. 3.3); variations with shower age (number of generations) for reasons of increased scattering at larger shower depths are only on the order of 10–20 % [27].

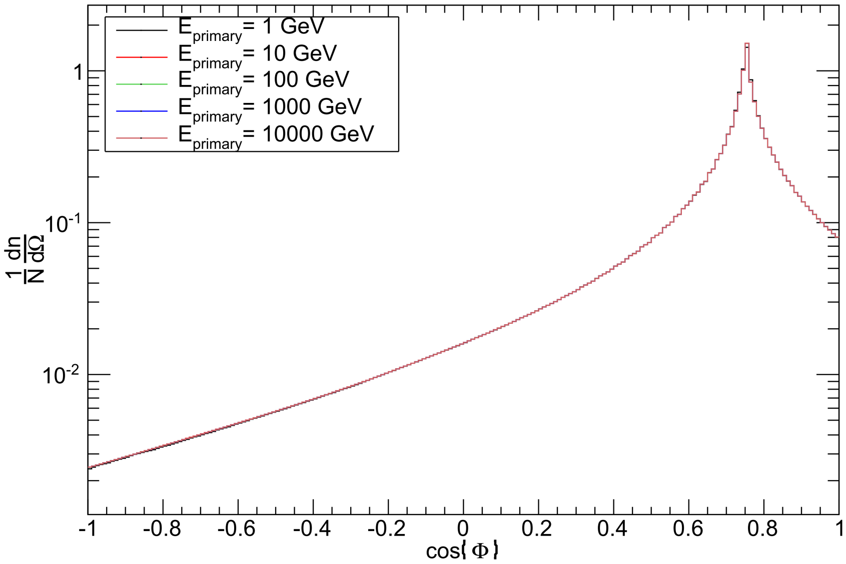

Figure 3.3.: Angular distribution (per photon and steradian) of Cherenkov photons emitted by electromagnetic showers, independent of primary energy. Figure taken from Ref. [27]

3.4.2. Hadronic cascades

Hadronic cascades can originate e.g. from neutral- or charged-current neutrino-nucleon scattering or from the decay of heavy particles. Characteristic length scales of longitudinal cascade development are governed by the nuclear interaction length, i.e. a hadron's mean path length before it interacts inelastically with a nucleus, $\lambda_I \approx 90.8\,\mathrm{cm}$ in ice [26].

Every hadronic cascade also comprises an electromagnetic component. The energy transfer to the latter proceeds via the production of π^0, which immediately decay into a pair of high-energy photons, thereby inducing overlaid electromagnetic cascades. Large shower-to-shower fluctuations in energy deposition are attributable in part to this process, and in part to the possibility that lower energy charged pions and other hadrons decay into muons, which are lost to further shower development [32]. As soon as the energy of a secondary drops below the pion-production threshold, it will no longer contribute to the energy transfer to the electromagnetic part of the cascade; the electromagnetic fraction rises with increasing energy of the primary hadron [34].

In general, hadronic cascades produce a smaller amount of Cherenkov light than do electromagnetic cascades. As a consequence, the energy deposit that is potentially detectable (also called *visible energy* [35]) by detectors sampling Cherenkov light is reduced. Causes for this suppression include the existence of slow neutrons, interactions leading to reduced secondary energies through binding energy losses, as well as increased Cherenkov energy thresholds (Eq. (3.6)) [36]. The quantity that governs the relative light yields is the ratio of the total (radiating) track lengths T_{had} and T_{em} [36],

$$F = T_{\mathrm{had}}/T_{\mathrm{em}} \ , \tag{3.13}$$

which has been simulated and fit as a function of energy in ice in Ref. [36]. On average, a 100 GeV hadronic cascade, for example, produces only around 75 % of the amount of Cherenkov light emitted by an electromagnetic cascade of the same energy, with event-to-event fluctuations of approximately 17 % with respect to the average.

3.5. The IceCube detector and its low-energy extension PINGU

3.5.1. Detector layout

PINGU (**P**recision **I**ceCube **N**ext **G**eneration **U**pgrade) [1] is a proposed low-energy extension to the existing *IceCube Neutrino Observatory*, constructed at the location of the Amundsen-Scott South Pole Station in Antarctica. The observatory consists of the cubic-kilometre *IceCube* array and two smaller subarrays, *DeepCore* and *IceTop*. The latter is a surface air shower array that can be used as a veto for backgrounds arising from extensive air showers and for studies of cosmic rays [37].

A hexagonal grid of 78 strings spaced approximately 125 m apart constitutes the IceCube array. Each string is instrumented with 60 optical sensors (DOMs, see Sec. 3.5.2 below) spaced 17 m apart vertically, at depths of 1450–2450 m below the surface [37].

The DeepCore subarray encompasses eight additional high-density strings and the seven central IceCube strings. Horizontal inter-string distances are 42 m and 72 m, respectively. The dedicated DeepCore strings are not instrumented in the region between 1850 m and 2100 m depths, where an elevated concentration of dust in the ice leads to significantly reduced scattering and absorption lengths (see Fig. 3.7). Directly above this region, there are 10 DOMs on each string with a DOM-to-DOM spacing of 10 m, while 50 DOMs with 7 m vertical spacing are located below. The majority of the DeepCore DOMs are equipped with photomultipliers whose quantum efficiency exceeds that of standard IceCube DOMs by around 35 %; in combination with the dense geometry and the option to use the upper DeepCore volume as well as surrounding strings as a veto against downward-going muons, this effectively reduces IceCube's detection threshold to around 10 GeV [38].

Conceptually following IceCube-DeepCore's hardware design, PINGU will be an infill within the lower region of the central DeepCore volume, where the ice is exceptionally clear. Extensive Monte Carlo optimisation studies have resulted in a baseline geometry (V36) consisting of 40 strings, each instrumented with 96 optical modules at depths between around 2130 m and 2415 m. The horizontal inter-string and vertical inter-DOM spacings are 22 m and 3 m, respectively. Thereby, PINGU will be able to effectively detect neutrinos with energies as low as a few GeV. Fig. 3.4 illustrates the scale of the PINGU detector in comparison with that of the surrounding DeepCore array. Optical modules employed by the PINGU detector will be updated and adapted versions of the existing IceCube-DeepCore DOM, the main components of which are described in Sec. 3.5.2 below. Refined optical module designs based on multiple PMTs or a combination of wavelength shifting and light guiding materials are currently under investigation [1].

3.5.2. Data acquisition

The Digital Optical Module

At the heart of the optical detection and data acquisition in the IceCube detector lies the *DOM* (**D**igital **O**ptical **M**odule): it incorporates a photomultiplier tube (PMT) and the DOM mainboard (MB) with its signal processing electronics within a spherical pressure-resistant glass housing. A flexible gel is used for optically coupling the PMT face to the surrounding glass sphere, simultaneously providing mechanical support. Together with the glass, it leads to the short wavelength cutoff of the DOM at around 350 nm. Since magnetic fields can significantly decrease the anode sensitivity of a PMT by reducing the efficiency with which photoelectrons are focussed onto the first dynode [39, Fig. 13-8], the IceCube PMT is shielded against Earth's magnetic field by a mu-metal grid.

Beside capturing and digitising the PMT anode signal, the DOM MB also provides an interface to the flasher board. The latter hosts a set of 12 LEDs able to produce bright ultraviolet optical pulses, which can be used to study properties of the Antarctic ice or to calibrate distant DOMs. In addition to this, the DOM MB drives another ultraviolet LED at 374 nm used for self-calibration purposes. Fig. 3.5 shows a diagram depicting the DOM and the arrangement of its main elements. Data transfer and power lines pass through the glass pressure shield via the electrical penetrator located at the top of the DOM.

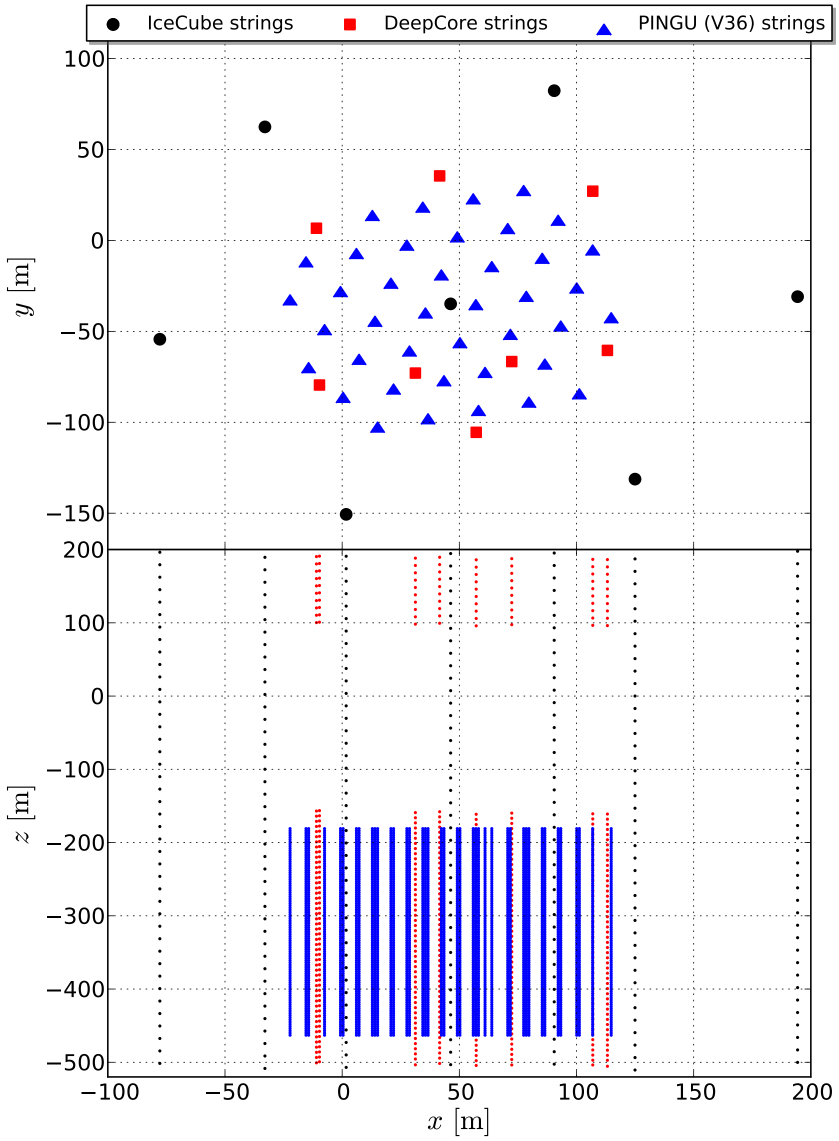

Figure 3.4.: Layouts of the existing DeepCore array and the perceived PINGU (V36) infill. The diagram at the top shows the string positions in the horizontal, given with respect to the origin of the detector coordinate system. Below, for depths at which DeepCore strings are instrumented, the vertical locations of optical sensors are shown; $z = 0$ corresponds to a depth of approximately 1948 m [40]. The high density of PINGU photocathodes (blue) is responsible for the low neutrino detection threshold of a few GeV. See text for details.

Upon detection of one or several photons, a DOM is to produce a digital output record, the “hit” [41]. Fig. 3.6 shows a block diagram of the DOM signal processing circuitry. At the output of the transformer coupling the PMT base to the DOM MB (top left), the analog PMT signal is split into three different signal paths. The topmost path allows the DOM MB to trigger on PMT signals. It hosts two comparators, one with a high resolution to sense PMT pulses initiated by single photoelectrons (SPE), and a reduced-sensitivity one for multiple-photoelectron (MPE) pulses. PMT trigger levels for in-ice DOMs are typically around 0.2 SPE (cf. Fig. 4.1).

Different *local coincidence* (LC) modes ultimately determine whether data capture is initiated. Upon triggering, each DOM transmits so-called LC tags to its nearest neighbours above and below. If these also exhibit discriminator threshold crossings within a time window of length 1 µs from the moment of tag reception, local coincidence is satisfied. Readouts resulting from local coincidence are called H(Hard-)LC readouts, while those without local coincidence are referred to as S(Soft-)LC readouts. Since it is much more likely that isolated hits leading to SLC readouts are the result of PMT dark noise rather than that of physics events, one can limit data flow by reducing the level of waveform detail for these types of hits. The capturing of the waveforms is described in the following.

ATWD From the DOM MB input, the main signal path sends the PMT signal through a 75 ns delay line to three pulse-shaping amplifiers of different gain (×16, ×2, ×0.25), which are each connected to a dedicated channel of two custom designed integrated circuits called *Analog Transient Waveform Digitiser* (ATWD). A discriminator threshold crossing triggers ATWD capture, resulting in the first 420 ns of the amplified PMT anode voltage being stored in analog memory at a rate of 300 MSPS (corresponding to a sampling period of 3.3 ns). If the local coincidence condition (HLC) is met, 128 10-bit analog-to-digital converters digitise the recorded voltages, which are then transferred to the FPGA. Starting with the highest-gain channel, if any one sample in the channel exceeds a certain number of counts corresponding to the saturation threshold, the next-highest gain channel is also digitised to ultimately ensure that the peak amplitude is recorded. Since a DOM might be triggered again during the 29 µs it takes the ATWD to digitise and transfer the input signal to the FPGA after capture, each DOM is equipped with a second ATWD to effectively reduce dead time to zero.

fADC For HLC hits, the third signal path accommodates for physics signals lasting longer than the 420 ns period recorded by the ATWD. After being routed through a set of pulse-shaping amplifiers, the signal is continuously sampled by a pipelined digitiser (fast ADC or fADC) at a rate of 40 MSPS, leading to a considerably longer sampling period of 25 ns compared to that of the ATWD. The total length of the waveform recorded by the fADC is 6.4 µs. Only three samples of the fADC record are saved in the case of an SLC hit. These are chosen such that a coarse charge stamp can be assigned and the time of the underlying SPE inferred, which both might provide useful information for global triggers.

Triggers Based on the global hit pattern, software running at the *IceCube Laboratory* (ICL) at the surface decides upon whether the whole detector is read out and an event is formed. Several distinct software triggers exist, which can also be applied to simulated data (chapter 4). The PINGU *Simple Multiplicity Trigger* (SMT3) is the most relevant for the studies presented in this work. It requires three close-neighbour DOMs in PINGU or in the bottom part of the IceCube/DeepCore strings depicted in Fig. 3.4 to record locally-coincident hits within a time window of 2.5 µs.

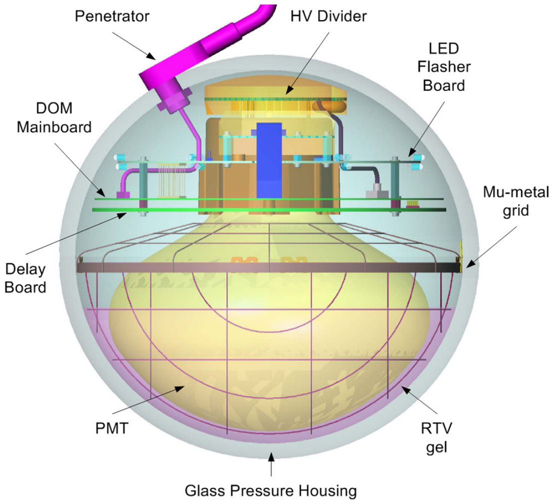

Figure 3.5.: The IceCube DOM comprises a 10-inch PMT and readout electronics arranged inside a pressure-resistant glass housing. For details see text. Illustration taken from Ref. [41].

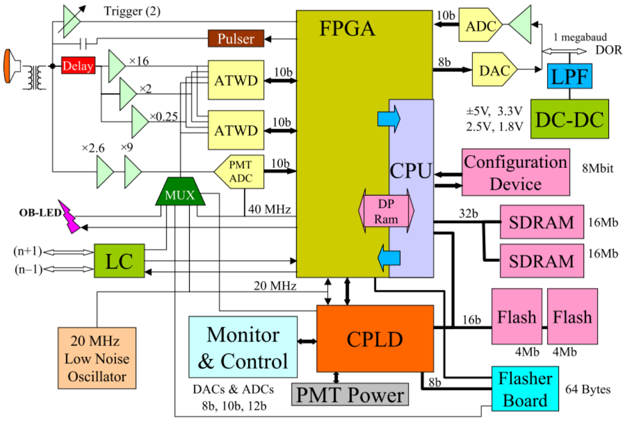

Figure 3.6.: The DOM mainboard holds all the signal processing circuitry necessary to generate a "hit". Details on the readout process are given in the text. Diagram taken from Ref. [41].

3.5.3. Photon propagation in Antarctic ice

The operation of IceCube depends crucially on the knowledge of the optical properties of the ice within and surrounding the detector. In order to be able to deduce trajectories and energy depositions of secondaries produced in neutrino-nucleon interactions in the detector vicinity from the Cherenkov photons registered by the optical sensors, scattering and absorption in the relevant optical and near ultraviolet wavelength regimes needs to be modelled accurately.

Some important concepts used in describing the propagation of light in deep ice come from Mie theory [42], which assumes scattering off of spherical particles. In deep ice, these particles are air bubbles with radii in the sub-mm range or dust particles which are some 10 to 100 times smaller. When air bubbles and dust particles occur at similar concentrations, the scattering rate between photons and dust particles can be neglected because the scattering length[1] $\lambda_\mathrm{s} \propto 1/n\langle\pi r^2\rangle$, where n is the number density of particles with average squared radius $\langle r^2\rangle$ [43]. The anisotropy of the scattering process is quantified by the average cosine of the scattering angle, $g \equiv \langle\cos\theta\rangle$, which takes into account properties of the angular distribution of the scattered photons. In the case of isotropy, $g = 0$, and one scatter is already sufficient to randomise the direction of a photon. A convenient measure of the *transport length* [44] in the case of non-isotropic scattering problems is the so-called *effective scattering length*

$$\lambda_\mathrm{e} = \frac{\lambda_\mathrm{s}}{1-g} \,, \tag{3.14}$$

with the *effective scattering coefficient* $b_e \equiv \lambda_\mathrm{e}^{-1}$ being its inverse. Similarly, one can define the *absorption coefficient* $a \equiv \lambda_\mathrm{a}^{-1}$, which consists of a component related to absorption by dust particles and a second component describing the intrinsic absorption by pure ice [45].

Efforts towards increasingly accurate determination of the properties of deep South Pole ice have been ongoing for more than two decades. First in-situ measurements [43] were performed at moderate depths between 800 m and 1000 m by the laser calibration system of AMANDA (Antarctic Muon And Neutrino Detector Array), at a wavelength of ∼ 515 nm. They found significantly broadened and delayed photon travel-time distributions between optical modules at different locations in the ice, consistent with scattering on a high residual density of air bubbles[2], and more importantly large absorption lengths, implying low concentrations of absorbing impurities, even in comparison to laboratory ice from purified water.

Follow-up measurements of both absorption and scattering extended to the 410–610 nm wavelength range [46] uncovered a rapid increase of absorption with wavelength as well as different absorption lengths at the top and the bottom of the detector. An observed decrease of the inverse scattering length with depth, independently of the wavelength, can be explained by the gradual phase transition of bubbles to solid clathrate hydrates at depths greater than approximately 800 m due to pressure increase [47]. Due to the refractive index of air-hydrate crystals exceeding that of pure ice by only around 0.4 %

---

[1] Average distance between two successive collisions.

[2] $g_\mathrm{bub} \approx 0.75 \Leftrightarrow \langle\theta_\mathrm{bub}\rangle \approx 41°$ [43]

[48], they are almost invisible in the ice and thus—in contrast to bubbles—only contribute weakly to the scattering of optical photons.

Data collected at depths extending to more than 2000 m lead to the conclusion that at ∼ 1400 m bubbles are no longer present, leaving the residual concentration of dust in the ice as the primary cause for depth-dependent structures exhibited by both the absorption and effective scattering coefficient. Simultaneously, different peaks in absorption and scattering as a function of depth could be correlated with corresponding peaks in the dust concentrations found in ice cores collected at other Antarctic locations, which allowed a first age vs. depth estimate [49].

Fig. 3.7 shows the effective scattering and absorption parameters at a wavelength of 400 nm determined from a global fit to in-situ light source data collected in 2008, using the IceCube LED calibration system [45] (“SpiceMie”). In addition, the figure shows the older AHA (Additionally Heterogeneous Absorption) ice model [50], which is an updated version of the ice description resulting from earlier measurements with flasher LEDs in AMANDA [28]. The four distinct peaks occurring between approximately 1500 m and 2100 m are the result of cold periods during the last glacial period in the late Pleistocene, with the most pronounced peak at ∼ 2000 m corresponding to an age of ∼ 65 000 years [28]. PINGU will be deployed below this “dust layer”, where the ice is clearest.

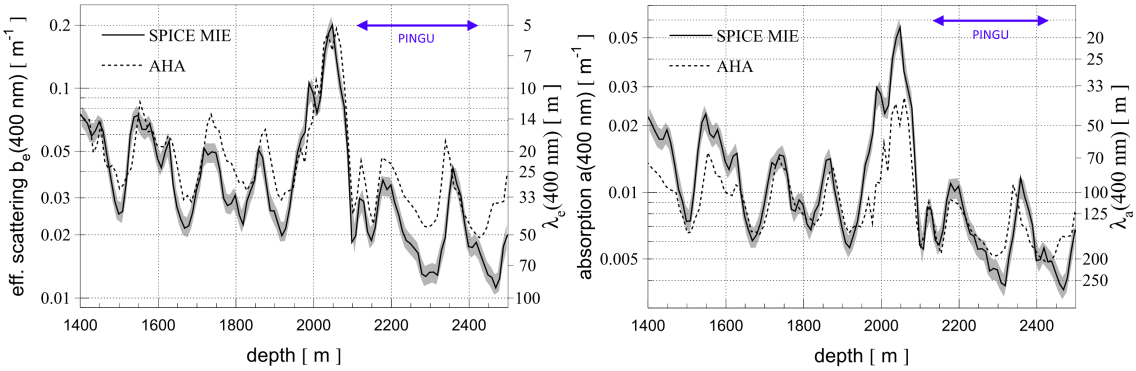

Figure 3.7.: Effective scattering and absorption parameters (coefficients and lengths) at λ = 400 nm as a function of depth in Antarctic ice, measured with AMANDA (“AHA” [50], dashed line) and IceCube (“SpiceMie”, solid line). See text for details. Figure taken from Ref. [45].

Finally, it should also be mentioned that a column of ice of approximately 30 cm in radius that has modified optical properties as the result of refreezing surrounds IceCube strings. It is referred to as “hole ice” and the increased scattering within is taken into account by an empirical model modifying the nominal angular sensitivity curve of an optical module [45]. For PINGU, improvements are planned in this respect [1].

4. Simulation and Low-Level Reconstruction

In order to reduce the risk of falling prey to bias, physics analyses in IceCube generally make strong use of Monte Carlo simulated data. Simulation includes the whole process from generating the primary particles of interest, modelling of interactions outside and within the detector, propagating the photons within the detector, to the DOM electronics and read-out. The different steps of the analysis can then be refined on the simulated data as well as on a small subset (10 %) of the experimental data in question, called the burn-sample. The approval of the collaboration's review committee in charge will then lead to the unblinding, in the course of which the analysis is expanded to include the whole data set.

On the other hand, since it provides the analyser with the true physics parameters of an event, simulation is essential for quantifying the performance of analysis and reconstruction methods.

The following will give a description of the tools currently used for simulation in PINGU. Most of the tools' implementations are also part of standard IceCube-DeepCore simulation.

4.1. Event generation

The generation of events taking place in the PINGU detector relies on the GENIE (**G**enerates **E**vents for **N**eutrino **I**nteraction **E**xperiments) Neutrino Monte Carlo generator [51]. GENIE currently takes into account the dominant scattering mechanisms in the energy range between several MeV and some hundred GeV, and can model interactions of any combination of neutrino flavour and target type. It generates an isotropic flux of neutrinos of the desired flavour following a user-defined power law distribution in energy on the surface of a predefined generation volume cylinder, upstream of the detector. A given neutrino is then forced to interact inside the generation volume, which contains the actual detector volume as a whole. Secondary particles from the interaction are fully simulated. The position of the interaction vertex along the cylinder axis is sampled from a uniform distribution over the cylinder length. GENIE calculates the true interaction probability, which is used in order to assign a statistical weight to the event in question. It is also possible to reweight the simulated energy spectrum to a physical one, e.g. the atmospheric neutrino flux spectrum.

4.2. Particle propagation

The output from GENIE ends up in a tree-like structure (`I3MCTree`) containing the primary (neutrino) and all of its daughter particles (secondaries), including their energies and directions at the point of interaction. While it is crucial to correctly calculate the light-yield of a given source throughout the whole detector volume, accurate closed form expressions do not exist due to to the heterogeneous (and anisotropic) nature of the Antarctic ice as well as the fact that the characteristic distance between a source and an observer (PMT) in the detector amounts to a few multiples of the scattering length [52]. As a result, any sophisticated simulation needs to be performed numerically.

The most recent IceCube software tool used for this purpose is `clsim` [53]. It can be used with parametrisations of cascade light yields, but also provides the option to perform a full *Geant4* simulation. In the case of the latter, the secondaries are propagated through the detector medium together with any further daughter particles produced, at least as long as their energies exceed the Cherenkov threshold. In this process, each Cherenkov emitter's trajectory is divided into a series of light-emitting segments, each of which is assumed to have a constant velocity $\beta = v/c$.

Subsequently, by exploiting the depth-dependent optical properties of the ice, i.e. scattering and absorption lengths, it propagates the photons created by the Cherenkov emitters one-by-one. Since `clsim` is based on the OpenCL framework, it can be configured to run in-parallel on GPUs or multi-core CPUs.[1]

Every photon that reaches a surface of a DOM carries information about its direction, wavelength, hit position and time, how often it has scattered, and properties at the time of emission, which are then used together with the DOM's acceptance and angular sensitivity to transform a photon into a photoelectron (`I3MCPE`), i.e. an electron ejected from the photocathode.

4.3. Detector simulation

The studies presented in this work are performed using simulated PINGU detector geometries consisting of DOMs with identical properties, since they are collectively modelled after a specific optical module within the DeepCore sub-array.

Noise and photomultiplier response

Prior to the simulation of the actual detector response, `Vuvuzela` [54] adds noise to the photoelectrons from the previous step. Apart from a thermal contribution of constant rate, it also generates a correlated component stemming from the decay of impurities and subsequent scintillation in the glass of the DOMs. After that, a dedicated module, the `PMTResponseSimulator` [55], converts the "raw" photoelectrons at the photocathode

---

[1] The more conventional way of modelling the arrival times of photons at the photo-detectors makes use of large look-up tables. Each photon table has to be generated in a time-consuming process in advance, and is only valid for that precise set of optical properties which constitutes the ice model at the basis of the table [53].

into a series of “weighted” photoelectrons or pulses, which are characterised by their arrival time and charge (a real-valued number). PMT jitter is simulated by shifting the times of `I3MCPE` initiating regular pulses (as opposed to pre-, late or afterpulses, see below) according to the distribution shown in Fig. 4.1 on the left, which has a Gaussian-equivalent width of approximately 2 ns. Note, however, that the time accompanying a `weightedI3MCPE` generated by this module does not include the PMT transit time it takes the “original” photoelectron to reach the first dynode and the emerging electron shower to traverse the PMT.

A regular pulse’s charge is distributed according to the sum of a Gaussian peaked at 1 photoelectron (PE) and a subdominant exponential component reflecting the existence of smaller amplitude pulses, which shift the mean charge to $\sim 0.85\,\mathrm{PE}$.[2] This distribution is shown on the right hand side of Fig. 4.1.

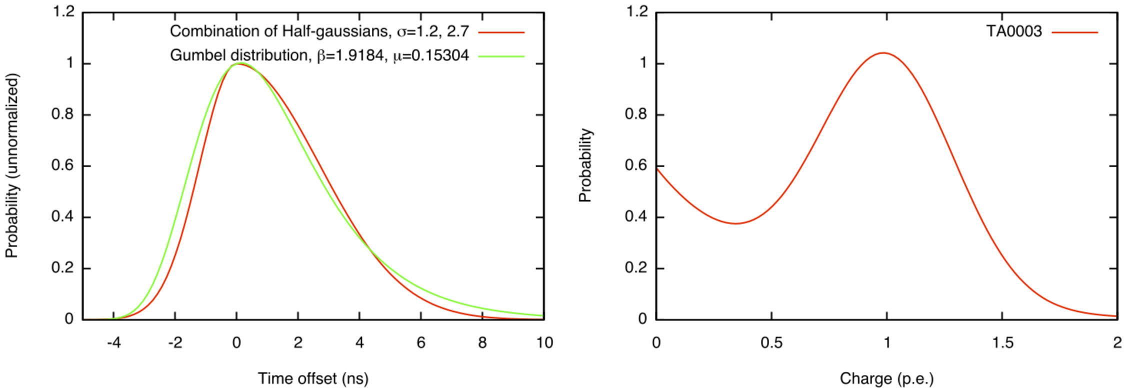

Figure 4.1.: *Left:* Single photoelectron (SPE) time distribution as used in the simulation of unscattered pulses in IceCube DOMs (green). The underlying Fisher-Tippett distribution can be found in [57]. For comparison, the sum of two half-gaussians is shown in red. *Right:* SPE charge response, consisting of a Gaussian peaked at 1 PE and an exponential component of smaller amplitude pulses. The small amplitude part (below the discriminator threshold) is important for events in which a DOM records several PEs. The fit can be found in [56, p.15, Eq. (1)].
Figures taken from Ref. [55].

In addition to the “standard” pulses described above, a PMT signal in rare cases may also stem from three other distinct sources, namely *prepulses*, *late pulses* and *afterpulses*. While prepulses occur on the order of 30 ns early, late and afterpulses can occur as early as regular pulses and up to ∼ 11 µs later, respectively [58, 59, 60].

In-depth explanations of the measurements after which all the physics effects described in this section are modelled can be found in Ref. [56].

---

[2] Dedicated measurements of a large amount of charge response distributions at different gains exhibit only moderate deviations, thus justifying the use of an average, gain-independent model in simulation as well as in reconstruction [56]. The physical charge can then be obtained by multiplying the pulse charge by the gain. At a gain of 10^7 the peak would be located at 1.6 pC.

DOM launching

`DOMLauncher` [61], the next module in the simulation chain, takes the pulses from `PMTResponseSimulator` arriving at the PMT anode and models their processing by the DOM mainboard electronics and logics units: a DOM is only triggered when the signal passes its discriminator threshold, and, depending on the type of local coincidence (see Sec. 3.5.2), the digitised ATWD or fADC waveforms are simulated—based on SPE pulse templates retrieved from calibration—and stored in the according `I3DOMLaunchSeries` object, taking into account e.g. the PMT transit time and digitiser saturation.

4.4. First reconstruction

Particle reconstructions algorithms usually do not handle the raw ATWD and fADC waveforms (Sec. 3.5.2) by themselves. Instead, they work with more physically motivated quantities, namely photon arrival times and pulse charges extracted from the waveforms. The general principles behind the two processing steps needed in between—*wave calibration* and *pulse extraction*—will be laid out in the following.

Wave calibration

"`I3WaveCalibrator` is a module that applies calibration constants to transform the contents of raw DOMLaunches (ADC counts) into calibrated waveforms (mV), while correcting for known effects of the electronics" [62].

The first effect that is corrected for is the digitiser baseline, i.e. the average digitiser count when the input voltage is zero. In order to obtain a voltage, the raw ADC counts are baseline-subtracted and multiplied by the gain [63]. In the second calibration step, the voltage-dependent time it takes a photomultiplied pulse to traverse the length of the PMT is subtracted from the time associated with a DOM launch, and an additional time offset between ATWD and fADC waveforms, arising from the delay line on the DOM MB, is corrected for. Furthermore, waveform distortions induced by the toroidal transformer AC coupling the high-voltage PMT anode to the front-end amplifiers on the DOM MB—the droop [64]—is taken into account, using an empirical model [65] whose parameters have been fit to measurements [66, pp. 6-9].

Pulse extraction

The extraction of a reconstructed number of photons per unit time from the waveforms—which then enable one to infer physical properties of the underlying event on the basis of a suitable description of the detector response—is handled by the `Wavedeform` [67] module, developed in Ref. [68]. It is motivated by the fact that several electronics components introduce characteristic shapings of the initially undistorted PMT anode waveform, meaning that the calibrated ATWD and fADC waveforms as such only have an "implicit" imprint of the true photon arrival time distribution.

Typical SPE waveforms as seen at the secondary of the PMT base transformer have a Gaussian peak of width $\sigma = 3.2\,\text{ns}$, yielding approximately 90 % of the total charge within a period of 10 ns following the maximum [56]. While a fraction of this width is due to the transformer, the amplifier stages in front of the different ATWD channels and the fADC apply further shaping functions of their own to the waveforms.[3] `Wavedeform` makes use of a *non-negative least squares* algorithm to reconstruct the waveforms, which are interpreted as linear superpositions of the SPE response function, and has proven to reach both high charge and time resolution [68]. Fig. 4.2 shows the result of an unfolding of a complex waveform.

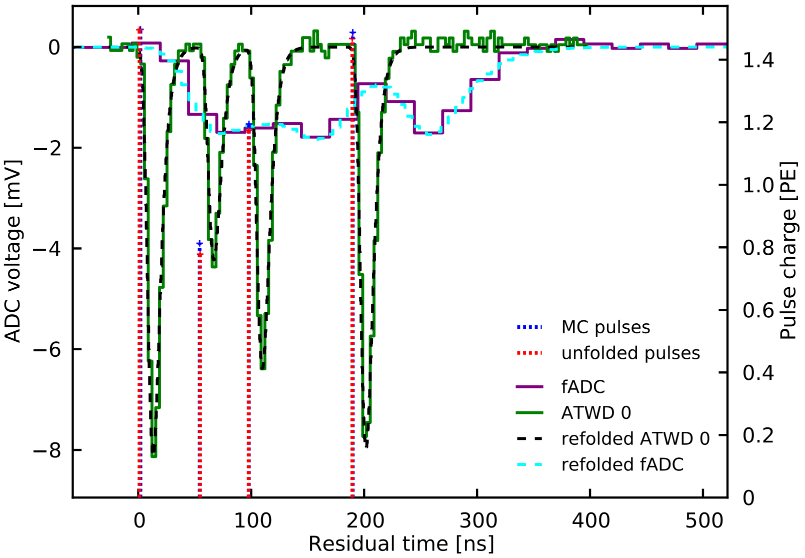

Figure 4.2.: Result of the unfolding of a complex waveform. The calibrated ATWD (green) and fADC (purple) readouts are the result of the four MC pulses shown in dark blue. One can see that the earliest photon arrives almost without any scatter, and leads to a regular pulse of ∼ 1.5 PE in magnitude. In this specific example, the second pulse, occurring approximately 50 ns later, is a late pulse brought about by inelastic backscattering of the photon at the first dynode.

Some reconstruction algorithms also require hits arising from noise to be removed. For an illustration of the principles behind the so-called *seeded RT-cleaning*, see e.g. Ref. [69].

---

[3] The corresponding characteristic shaping times significantly exceed the typical 3.2 ns SPE waveform width. Shaping functions for the highest-gain ATWD channel and the fADC are depicted in Ref. [68, Fig. 3.6].

5. Parameter Estimation and Likelihood Concepts

A measurement in physics is often subject to randomness, in the sense that the outcome of its repeated performance cannot be predicted. The concept of probability quantifies the degree of randomness, whether the latter is due to errors related to the measurement process or whether it is the consequence of an inherent (e.g. quantum mechanical) unpredictability [70].

The same concept of probability affects event reconstruction in IceCube: starting off with a measurement, i.e. the detector readout, and with a model relating the readout with the underlying event, the aim of reconstruction is to make inferences about the true nature of the event.

This chapter will first introduce Bayes' theorem, which is frequently employed in data analysis, before turning to the general concept of parameter estimation and presenting some useful properties of estimators. It will conclude with the method of maximum likelihood—used in the studies presented in this work—and will show how the statistical uncertainty of estimators obtained in this way can be quantified.

5.1. Bayes' theorem

Bayes' theorem states that the conditional probability "A given B"

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \ . \tag{5.1}$$

From the Bayesian point of view, the set S consisting of subsets $A, B, \ldots$ is a hypothesis space: it is made up of individual statements that are either true or false. The probability assigned to the hypothesis A, $P(A)$, quantifies one's degree of belief about whether A is true.

In the frequentist scenario, the set S consists of all the possible outcomes (events) of an experimental trial or measurement, which is thought of as being repeatable (even if only hypothetically). A number of elementary outcomes constitute each subset A.

In data analysis, replacing A and B in Bayes' theorem by *hypothesis* and *data* permits one to quantify the probability that the hypothesis is true given the data observed [71]:

$$P(\text{hypothesis}|\text{data}) \propto P(\text{data}|\text{hypothesis})P(\text{hypothesis}) \ . \tag{5.2}$$

The *prior* probability $P(\text{hypothesis})$ corresponds to one's initial degree of belief about the truthfulness of the hypothesis. This is subsequently adjusted by the *likelihood* function

$P(\text{data}|\text{hypothesis})$, which takes into account the measured data, to yield the *posterior* probability $P(\text{hypothesis}|\text{data})$. Omitting the normalising *evidence* $P(\text{data})$ in Eq. (5.2) is justified in the case of parameter estimation, as it does not explicitly depend on the hypothesis. Here, the hypothesis might represent specific numerical values of a set of parameters. In the continuous case, *probability* has to be replaced by *probability density function* (p.d.f.).

When one is ignorant, one might consider choosing a uniform prior, i.e. $P(\text{hypothesis}) =$ const. After including this constant term in the normalisation factor, Eq. (5.2) becomes a simple proportionality between posterior and likelihood:

$$P(\text{hypothesis}|\text{data}) \propto P(\text{data}|\text{hypothesis}) \; . \tag{5.3}$$

5.2. Maximum likelihood estimation

Assuming one has performed an experiment consisting of a set of n measurements $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ of a random variable x described by the p.d.f. $f(x|\boldsymbol{\theta})$, depending on m unknown parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_m)$, then the aim of parameter estimation is constructing one (or several) functions of the data x_k to infer the value(s) of the unknown parameter(s). Each such function is called an *estimator* of θ_i (denoted by $\hat{\theta}_i$). The quality of a given estimator can be gauged by evaluating its *consistency*, *bias* and *mean squared error*, which are defined as follows [70].

Consistency An estimator is called consistent if it converges (probabilistically) to the true value in the limit of many measurements n:

$$\lim_{n\to\infty} P(|\hat{\theta} - \theta| > \varepsilon) = 0 \;\forall\, \varepsilon > 0 \; . \tag{5.4}$$

Bias Being a function of the data, every estimator $\hat{\theta}(\boldsymbol{x})$ in itself is a random variable. Therefore one can define its expectation value $\mathrm{E}[\hat{\theta}(\boldsymbol{x})]$, which in turn determines the estimator's bias

$$b = \mathrm{E}[\hat{\theta}] - \theta \; . \tag{5.5}$$

An unbiased estimator has $b = 0$ independent of n, while $\lim_{n\to\infty} b = 0$ would correspond to an asymptotically unbiased estimator.[1]

Mean squared error The mean squared error of an estimator is defined as

$$\mathrm{MSE} = \mathrm{E}[(\hat{\theta} - \theta)^2] = \mathrm{V}[\hat{\theta}] + b^2 \; , \tag{5.6}$$

and hence can be interpreted as the superposition of a statistical and a systematic contribution.

---

[1] Since the expectation value is taken over an infinite number of experiments, consistency of an estimator only implies that its bias vanishes asymptotically. An unbiased estimator, on the other hand, is also consistent.

The "best estimate" for the hypothesis in Eq. (5.2) is the one for which the posterior has a maximum. Under the assumption of a uniform prior (Eq. (5.3)), this estimate is equivalent to the hypothesis that maximises the likelihood function, i.e. the Bayesian estimator corresponds to the *maximum likelihood* (ML) estimator.

When the n data $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ obtained from the experiment are independent, the likelihood function is given by the product of the individual probabilities [70]:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta}) \ , \tag{5.7}$$

where L is assumed to be a function of the unknown parameters $\boldsymbol{\theta}$ only (since the experimental results are fixed). The ML estimators $\hat{\theta}_i$ are then the solutions to the m simultaneous equations

$$\left.\frac{\partial L}{\partial \theta_i}\right|_{\theta_i=\hat{\theta}_i} = 0 \ , \quad i = 1, ..., m \ , \tag{5.8}$$

given that the second derivatives with respect to the same parameters are negative. In practice, however, one usually instead constructs the "log-likelihood function" [72]

$$l(\boldsymbol{\theta}) \equiv \ln L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln f(x_i|\boldsymbol{\theta}) \ , \tag{5.9}$$

which clearly has the same maxima as Eq. (5.7).

Asymptotic confidence regions

The expansion of the log-likelihood function (5.9) in a series about the point $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_m)$ yields

$$l(\boldsymbol{\theta}) = l(\hat{\boldsymbol{\theta}}) + \sum_{i=1}^{m} \left(\frac{\partial l}{\partial \theta_i}\right)_{\hat{\boldsymbol{\theta}}} (\theta_i - \hat{\theta}_i) + \frac{1}{2} \sum_{j=1}^{m} \sum_{k=1}^{m} \left(\frac{\partial^2 l}{\partial \theta_j \theta_k}\right)_{\hat{\boldsymbol{\theta}}} (\theta_j - \hat{\theta}_j)(\theta_k - \hat{\theta}_k) + \cdots \ , \tag{5.10}$$

which, from the definition (5.8) of $\hat{\boldsymbol{\theta}}$, simplifies to

$$l(\hat{\boldsymbol{\theta}}) - l(\boldsymbol{\theta}) = -\frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T A(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) + \cdots \ , \tag{5.11}$$

where $-A = \left(\frac{\partial^2 l}{\partial \theta_i \theta_j}(\hat{\boldsymbol{\theta}})\right)_{i,j=1,\cdots,m}$ is the matrix of second derivatives of l, evaluated at the position of the ML estimate $\hat{\boldsymbol{\theta}}$. In the asymptotic limit, i.e. for a large sample size $n \to \infty$, the elements of the matrix $-A$, which still depend on the observed sample, can be replaced by their expectation values, so that [72]

$$B \equiv E(A) = -\left(E\left[\frac{\partial^2 l}{\partial \theta_i \theta_j}(\hat{\boldsymbol{\theta}})\right]\right)_{i,j=1,\cdots,m} \ . \tag{5.12}$$

By exponentiating Eq. (5.11) and neglecting terms of higher order, the likelihood function becomes

$$L(\boldsymbol{\theta}) = L(\hat{\boldsymbol{\theta}}) \exp\{-\frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T B(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})\} \ , \tag{5.13}$$

i.e. an m-dimensional normal distribution with mean $\hat{\boldsymbol{\theta}}$ and *covariance matrix* $C = B^{-1}$. The diagonal elements $c_{ii} = \sigma^2(\hat{\theta}_i)$ are the variances of the ML estimators $\hat{\theta}_1, \cdots, \hat{\theta}_m$, while the off-diagonal elements of C are given by the covariances $\mathrm{cov}(\hat{\theta}_i, \hat{\theta}_j) = \rho\sigma(\hat{\theta}_i)\sigma(\hat{\theta}_j)$.

In the asymptotic limit, it is true in general that both the joint p.d.f. $g(\hat{\boldsymbol{\theta}})$ for the estimator $\hat{\boldsymbol{\theta}}$ as well as the likelihood function $L(\boldsymbol{\theta})$ become Gaussian, centered about the true and the estimated parameter values, respectively [70]. Contours of constant probability density or likelihood are m-dimensional hypersurfaces and correspond to constant $Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = Q(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T C^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})$. Q is distributed according to a $\chi^2(m)$ distribution, with the number of estimated parameters m corresponding to the number of degrees of freedom (d.o.f.) [70].

One can make a statement about the probability β that the likelihood of the whole set of true parameter values lies within a certain region, i.e.

$$P\left[Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \leq K_\beta^2\right] = \beta, \tag{5.14}$$

where $\int_{-\infty}^{K_\beta^2} \chi^2(Q; m)\, \mathrm{d}Q = \beta$. Fig. 5.1 shows the contour of constant likelihood defined by $K_\beta^2 = 1$ in the two-dimensional case.

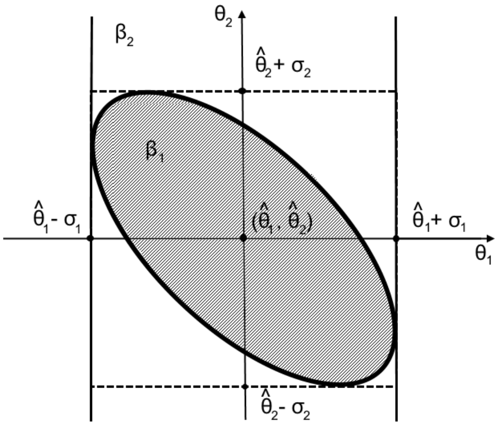

Figure 5.1.: Asymptotic confidence region for the true parameter values (θ_1, θ_2) for $K_\beta = 1$ and correlation coefficient $\rho = -0.5$. The probability content of the inner ellipse $\beta_1 = 39.3\,\%$. The vertical strip $\beta_2 = P\left[\hat{\theta}_1 - \sigma_1 \leq \theta_1 \leq \hat{\theta}_1 + \sigma_1\right] = 68.3\,\%$ defines a one-dimensional confidence interval for the true parameter value θ_1, marginalised over all possible values of θ_2 [73].

6. Event Reconstruction in PINGU

6.1. Cascade, track and hybrid event hypotheses

When setting up a hypothesis for cascade-like events, one typically assumes a seven-dimensional set of parameters: the vertex $\vec{x}_c = x_c\vec{e}_x + y_c\vec{e}_y + z_c\vec{e}_z$ and its corresponding time t_c, the direction (θ_c, ϕ_c) specified by zenith θ_c and azimuth ϕ_c, and finally a visible energy E_c.

The vertex coordinates (x_c, y_c, z_c) are taken relative to the origin of a cartesian coordinate system centered almost at the true geometric centre of the detector, while the two angles (θ_c, ϕ_c) characterise the radial vector pointing back toward the source of the incident particle which generated the cascade. This means that a vertically down-going particle has a zenith angle of 0, and a vertically up-going particle has one of π. Positions are given in metres, angles in radians.

When considering deep-inelastic ν_e-nucleon scattering at energies as low as the GeV range relevant to the measurement of the neutrino mass hierarchy, the energy deposited in Cherenkov-radiating charged particles in general cannot be considered as a good proxy of the true neutrino energy, because of the relative Cherenkov light yield suppression of the hadronic cascade (cf. Sec. 3.4.2). Since it is not possible to disentangle the hadronic and electromagnetic cascades even at the PINGU scale, one cannot simply deduce the energy of the neutrino from the reconstructed visible energy.

For this reason, a distinct reconstruction strategy is employed in PINGU: the cascade hypothesis described above is refined to also include the length of the daughter muon (the *track*, only truly existing in the case of ν_μ CC events) and optionally a second set of zenith and azimuthal angles to take into account non-collinearity of the hadronic cascade and the (hypothetical) muon [1]. A proxy of the true neutrino energy can then be obtained by scaling up the electromagnetic-equivalent energy of the hadronic cascade and adding it to the track energy,

$$\hat{E}_\nu = \hat{E}_{\text{casc}}/h_{\text{elm}\to\text{had}}(\hat{E}_{\text{casc}}) + \hat{E}_{\text{track}} \ . \tag{6.1}$$

Here, $h_{\text{elm}\to\text{had}}(\hat{E}_{\text{casc}})$ is the relative Cherenkov amplitude suppression of the hadronic cascade. Since a muon in this energy range can be assumed to be minimum-ionising (cf. Sec. 3.2.1), the conversion from track length to energy follows straightforwardly by applying a constant energy loss rate assumed to be $\Delta E/\Delta L \approx 1\,\text{GeV}/4.5\,\text{m}$. It turns out that Eq. (6.1) can provide a good estimate of the true neutrino energy even when the underlying event is a CC interaction of an incident ν_e (Ref. [1] and Sec. 8.3). A description of the reconstruction of a source's event parameters based on a maximum likelihood approach is provided in the following.

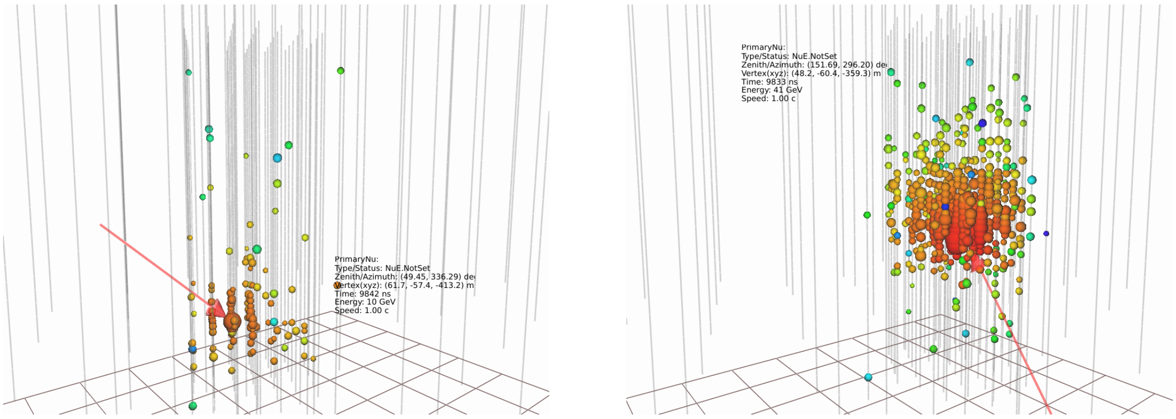

Figure 6.1.: Illustration of two simulated ν_e undergoing CC mediated deep-inelastic scattering interactions within the PINGU fiducial volume. The locations of the spheres show which DOMs have registered a signal during the detector readout time. Their sizes indicate the amount of accumulated charge, whereas the colour represents the pulse times, with red corresponding to earlier and blue to later times. In the left panel one can see a downward-going neutrino of 10 GeV energy, while in the right panel an upward-going neutrino deposits significantly more energy in Cherenkov radiating secondaries, thereby inducing a typical spherical cascade signature.

6.2. Maximum likelihood reconstruction

6.2.1. Poisson likelihood

When a cascade with an electromagnetic energy equivalent of E is contained within the detector, as is illustrated in Fig. 6.1 in an exemplary manner for ν_e CC interactions in PINGU, the number of photons k detected by any given DOM is expected to be Poisson-distributed around the mean μ. It is then the very regular nature of the cascade's light emission profile (cf. Fig. 3.3) as well as the linearity between energy and Cherenkov amplitude which allow one to write

$$\mu = \Lambda E \ , \tag{6.2}$$

where Λ is the number of photons per unit source energy arriving at the optical module. Alternatively, $\Lambda \cdot 1$ GeV is the number of photons produced by a "reference" cascade of 1 GeV energy [52]. The template function Λ is dependent on the source-observer configuration and on the ice properties, but not on the source energy, and it is evaluated from tabulated Monte Carlo simulation. In conjunction with the included normalised photon arrival time distributions, these tables provide the mean number of photons expected to be observed at any given position in the detector in any given time range (see Sec. 6.2.2).

A likelihood function, assuming fluctuations of purely Poissonian origin in collected charge, for a source with energy E resulting in the detection of $n_{D,i}$ photons in the DOM-time bin tuple (D, i) can be constructed as the product of the individual Poisson probabilities,

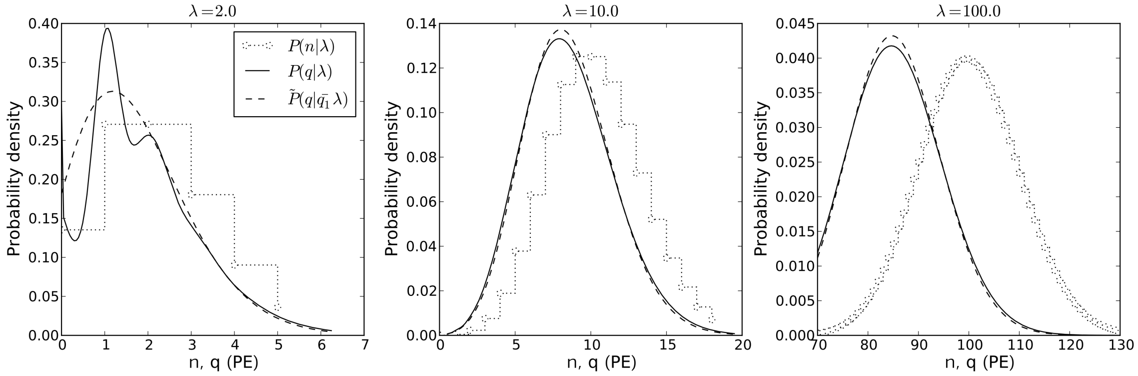

Figure 6.2.: In reconstruction, the actual PMT charge distribution (solid line), which corresponds to the numerical convolution of the Poisson distribution $P(n|\lambda)$ (stepped) with the PMT charge response function, is approximated by extending the Poisson distribution to real numbers (dashed line). The mean $\bar{q}_1 = 0.85$ of the SPE charge response function is responsible for the offset. Figure taken from Ref. [52].

$$L = \prod_{\substack{\text{DOMs}\\ \text{D}}} \prod_{\substack{\text{time bins}\\ \text{i}}} P_p(n_{D,i}|\mu_{D,i}) \equiv \prod_{\substack{\text{DOMs}\\ \text{D}}} \prod_{\substack{\text{time bins}\\ \text{i}}} \frac{\mu_{D,i}^{n_{D,i}}}{n_{D,i}!} e^{-\mu_{D,i}} \ , \tag{6.3}$$

where the mean amplitudes $\mu_{D,i}$ are obtained by multiplying Eq. (6.2) with the integrated probability (or *charge quantile*, see Sec. 6.2.2) that a photon emitted from the source in question arrives at DOM D during the interval spanned by the time bin i.[1]

For a given expected number of photons μ arriving at a DOM, however, the actual charge distribution $P(q|\mu)$ is obtained by convolving the Poisson distribution $P_p(n|\mu)$ (with integer n) with the PMT charge response function $P_c(q|n)$ (with real q), $P(q|\mu) = \sum_{n=0}^{\infty} P_c(q|n)P_p(n|\mu)$ [74]. It was found to be sufficient for both small and large mean amplitudes μ to analytically approximate $P(q|\mu)$ by extending the Poisson probability distribution to real numbers, i.e. replacing the normalising factorial $n!$ in Eq. (6.3) with the gamma function $\Gamma(n+1)$ [52, Fig. 10]. A comparison of the three distributions is shown in Fig. 6.2.

For a more general reconstruction scenario consisting of additional light sources and a PMT noise contribution, one substitutes

$$\mu \to \sum_{\substack{\text{sources}\\ \text{s}}} \Lambda_s E_s + \mu_{\text{noise}} \ . \tag{6.4}$$

The module which implements the logarithm of the likelihood (6.3) within the IceCube software framework, and which is employed in the most sophisticated PINGU event reconstruction algorithm, is called `Millipede` [75]. Given pulses extracted from the waveforms, one can specify whether DOMs without any reconstructed charge should be included in

---

[1] Implicitly, $i \equiv i_D$. The subscript is omitted to avoid (more) clutter.

the likelihood evaluation, whether the photon arrival times should be considered (and if so, the binning method), as well as a relative DOM efficiency correction to account for possible shadowing by signal cables. With photon timing applied, time ranges during which the waveform exhibits saturation can either be masked out, or the affected DOM can be excluded completely from the likelihood evaluation. DOM-specific quantities such as coordinates, relative efficiencies or (constant) noise rates are obtained from the detector's `GCD` file. For PINGU DOMs that are part of the simulated geometries used in this work, the dark noise rate is taken as $\sim 715\,\mathrm{Hz}$.

When used with its default binning scheme, for computational reasons `Millipede` only creates one single time bin spanning the whole readout time for DOMs which have not registered a signal. Even for hit DOMs, all bins that do not contain any charge and that lie between two signals are merged into one larger bin.

Fig. 6.3 illustrates how the full likelihood is obtained by summing over the contributions of all DOMs in the detector. Here, the underlying event is a low-enery ν_e depositing only around $1\,\mathrm{GeV}$ of energy in charged particles in a deep-inelastic CC scattering interaction, with a mere eight optical modules at distances between $30\,\mathrm{m}$ and $150\,\mathrm{m}$ from the true interaction point having registered at least one hit. The figure shows the likelihood as a function of the electromagnetic-equivalent energy for a single-cascade hypothesis after a maximum likelihood reconstruction has been performed with Multinest (see Sec. 6.2.3 below), with all the other parameter values fixed at their reconstructed values. One should note especially how the inclusion of unhit DOMs places tighter constraints on the position of the likelihood maximum.

6.2.2. Photospline tables

The fact that analytic functions for both photon time and amplitude distributions in the Antarctic ice sheet can at best provide approximate solutions [52, p.6] is the result of the scattering and absorption lengths for optical photons being depth-dependent and furthermore on the same order; the more sophisticated reconstructions rely on light yields and arrival times evaluated by Monte Carlo simulation of photon propagation in the ice, which are tabulated and subsequently fit by a multi-dimensional spline surface.

For ice models neglecting anisotropy and horizontal inhomogeneities, for a nearly point-like electromagnetic shower the number of necessary table dimensions is found to be six; the scaling function Λ in Eq. (6.2) is parametrised by the depth (z-coordinate) and zenith angle of the source, its relative spatial displacement from the observer, and the difference in time between the moment of photon emission and eventual detection. Light is produced according to shower parametrisations given in Ref. [32] and then propagated by the software package `PHOTONICS` [76], taking into account the scattering and absorption of a given ice model. Since the evaluation of the likelihood in reconstruction requires the Poissonian mean, each table bin entry is created from a large sample of the photon distribution in the bin. The final photon yield for a specific source-observer pair stored in a table has the wavelength acceptance and angular sensitivity of the optical module folded in.

So that the tables can be loaded into memory of ordinary computers without the introduction of binning artifacts distorting the expected amplitudes, very finely binned "raw"

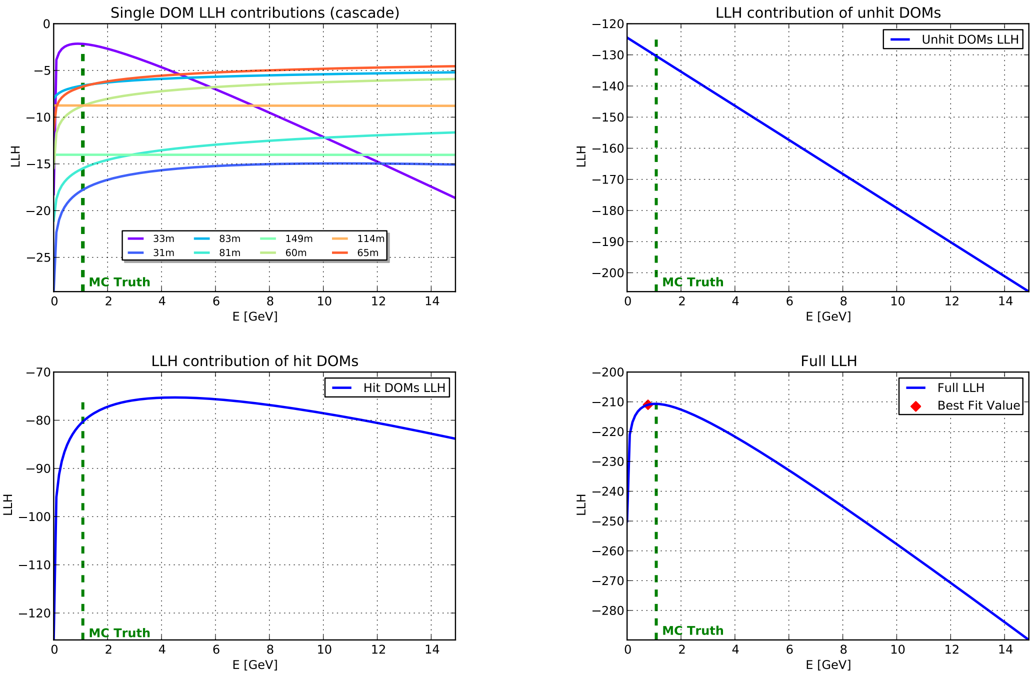

Figure 6.3.: Log-likelihood (LLH) scans in energy for a single-cascade hypothesis, with all other parameters fixed at their ML estimates. The full log-likelihood follows as the sum of the individual contributions. See text for details.

tables are produced and then processed by Photospline [77], which represents the light yields as a multi-dimensional spline surface. The latter has the additional advantage of being able to provide analytically computed gradients of both total amplitudes and arrival time quantiles, which will be drawn upon in Sec. 6.3.2. Moreover, it is possible to convolve the normalised photon delay time distributions with Gaussians of variable width, thereby incorporating the finite transit time spread of the PMT pulses into the modelling of the detector response.

While hit times in data are usually given in relation to the trigger range capturing the underlying event, photon arrival time quantiles and probability densities stored in spline tables are given with respect to some arbitrary "zero-point" in time. If reconstruction is supposed to make use of timing information, the bin-wise expectation values need to be evaluated correctly. This is obtained by setting the zero-point to the time an unscattered photon would have reached DOM D in question. For a point-like cascade, there is no delay between production of the primary particle initiating the cascade and photon emission, so the *direct time* is given by the sum of a time t_{casc} ascribed to the cascade-like source,[2] and the purely geometric propagation time of a photon that is not subjected to scattering,

---

[2] In the case of a deep-inelastic charged-current ν_e scattering, t_{casc} would describe the time of interaction as well as the emission time of all Cherenkov photons arriving at the sensor and originating from the interaction (electromagnetic and hadronic shower).

$t_{\text{geo}} = |\vec{x}_D - \vec{x}_{\text{casc}}|/c_{\text{group}}$, where $c_{\text{group}} = c_0/n_{\text{ice}}$ is the group velocity of light in ice.

The *residual time* t_{res} is then defined as a photon's time delay with respect to unscattered propagation,

$$t_{\text{res}} \equiv t - t_{\text{geo}} - t_{\text{casc}} \ . \tag{6.5}$$

When μ_D is the total charge expected to be registered by a DOM, and $P(t)$ is the cumulative photon arrival time distribution, the fraction which is expected in a time bin i with edges $t_2 > t_1$ is given by $q_i \equiv P(t_2) - P(t_1)$, so that

$$\mu_{D,i} = \mu_D \cdot q_i \ . \tag{6.6}$$

Since no photons are expected to arrive prior to the direct time, $P(t < 0) \equiv 0$.

Effects that are not included in the spline tables, but which should be accounted for before comparing the amplitudes with data or realistic simulation, are the fraction of PMT surface unshadowed by cable, the relative quantum efficiency of the photocathode and the mean of the SPE charge response. In Fig. 6.4 one can compare the tabulated photon delay time distributions and mean amplitudes for observers at different distances from a cascade source to results obtained from numerical simulations as laid out in Sec. 4.2. One can see that there are marked discrepancies in the total amplitudes but also in the delay time distributions.

6.2.3. Multinest

The Multinest algorithm [78] is a Bayesian inference tool for efficiently evaluating the Bayesian evidence introduced in Sec. 5.1. While the evidence itself does not play a role in parameter estimation problems as is the one of event reconstruction, posterior inferences arise simply because Multinest samples the whole likelihood space. The mechanics of the sampling process, starting from transforming the multi-dimensional evidence integral $Z = \int L(\boldsymbol{\theta})\pi(\boldsymbol{\theta})\mathrm{d}^D\boldsymbol{\theta}$, where $L(\boldsymbol{\theta})$ and $\pi(\boldsymbol{\theta})$ are the likelihood function and the prior, respectively, and D is the dimensionality of the likelihood space, into a weighted sum of likelihoods corresponding to increasing fractions of traversed prior volume, to approximating iso-likelihood contours in order to efficiently draw points with increasing likelihood and simultaneously identifying different modes of the posterior distribution, are discussed extensively in Refs. [78, 79] and references therein.

Furthermore, Multinest is the name of the reconstruction strategy [80] applied to neutrinos in the GeV range detected by PINGU. It makes use of the algorithm presented above in order to maximise the likelihood defined in Eq. (6.3) with respect to the parameters that constitute the hybrid event hypothesis described in the beginning of this chapter. The track part of the hypothesis is actually split up into several individual minimum ionising track segments of defined length L (and therefore defined energy $E = L\Delta E/\Delta L$), for which one uses adequate spline tables much as in the case of a cascade-like source to look up and scale the expected light yields of a template muon.

6.3. Semi-analytic estimation of reconstruction residuals

Sec. 5.2 demonstrated how the expectation value of the Hessian matrix of the log-likelihood function can be used to estimate the variances of the ML estimators. Since spline tables implement gradients for both total amplitudes and time-delay quantiles, it is possible to calculate gradients of the likelihood function itself. This functionality can be exploited by gradient-based minimisers in an ML reconstruction, but it also enables one to obtain the variances of ML estimators of the Poissonian likelihood (6.3) on an event-by-event basis.

6.3.1. Cascade resolution estimator

The recently developed `Resca` (**R**esolution **E**stimator for **C**ascades) [81] is an approach that makes use of the relation

$$\mathrm{cov}^{-1}(\theta_j, \theta_k) = \sum_{\mathrm{D,i}} \frac{1}{\mu_{D,i}} \frac{\partial \mu_{D,i}}{\partial \theta_j} \frac{\partial \mu_{D,i}}{\partial \theta_k} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{6.7}$$

in the asymptotic limit, where the sum ranges over all DOMs and time bins which are not excluded from the likelihood evaluation. The parameters θ_j are those that constitute the cascade hypothesis.

With Eq. (6.6), the partial derivatives of the mean amplitudes $\mu_{D,i}$ can be written as

$$\frac{\partial \mu_{D,i}}{\partial \theta_j} = \frac{\partial q_i}{\partial \theta_j} \cdot \mu_D + q_i \cdot \frac{\partial \mu_D}{\partial \theta_j} \ . \tag{6.8}$$

On the one hand, the second summand is identically zero when the parameter θ_j in question is the time, due to all total amplitudes being time-independent. When considering the source energy, on the other hand, the first summand vanishes, while the second summand is simply the product of charge quantile and light scaling factor $q_i \cdot \Lambda_D$ (using Eq. (6.2)).

In implementing Eq. (6.7), one can draw upon the response and gradient matrices constructed by `Millipede` after it has been provided with the cascade hypothesis. For only one source, the former can be regarded as an N-dimensional column vector, with N being the total number of time bins; its i-th component multiplied by the energy of the source corresponds to the expected charge in bin i, Eqs. (6.2) and (6.6). These entries also yield the partial derivatives (6.8) of the bin-wise amplitudes with respect to the source energy. The gradient matrix is of $6N$ shape, its elements also remain to be scaled by the source energy to obtain the partial derivatives (6.8) with respect to all source parameters apart from energy.

6.3.2. Extension to hybrid reconstruction

It would be desirable to extend the single-cascade approach from above to the multiple-source scenario used in reconstructing low-energy events in PINGU. Here, one does not only have to take into account global hypothesis parameters that might only describe the cascade (e.g. cascade energy or direction) or the track (e.g. track length/energy or direction), or both (e.g. vertex position, time or direction in the collinear case), but one also has to distinguish between those that apply to the track as a whole and those that apply to the individual track segments only (e.g. track segment vertex position, time).

If the aim is computing the covariance matrix of all eight or ten global hypothesis parameters, then the partial derivatives of the mean amplitudes with respect to the shared parameters $\theta_j = x, y, z, t$ can be written as

$$\frac{\partial \mu_{D,i}}{\partial \theta_j} = \frac{\partial \mu^{\text{casc}}_{D,i}}{\partial \theta_j} + \sum_{\substack{\text{track segments} \\ \text{s}}} \frac{\partial \mu^{\text{s}}_{D,i}}{\partial \theta_j} \,, \tag{6.9}$$

since an infinitesimal change in any of these global variables translates to the same change in the individual cascade and track segment parameters. Naturally, when the parameter in question is the cascade energy, the sum over the track segments s should be zero. Conversely, the cascade term should vanish when Eq. (6.9) is evaluated with respect to track energy; a single contribution is assumed to originate from the last track segment only, whose length is the difference between the total track length and the next-largest integral multiple of the track segment length. Finally, one should note that this approach can only provide an approximate expression of the change in amplitudes upon a change in the direction of the track. It will be assumed that an infinitesimal rotation in either zenith or azimuth is equivalent to rotating all of the track's individual segments through the same angle.

As opposed to the single source case, the `Millipede` response matrix is now of $N \times m$ shape, with N denoting the total number of time bins again, and m the number of sources; its (i, j)-th component multiplied by the energy of the source j corresponds to the expected charge in bin i from that source only. The gradient matrix is of $m \times 6N$ shape.

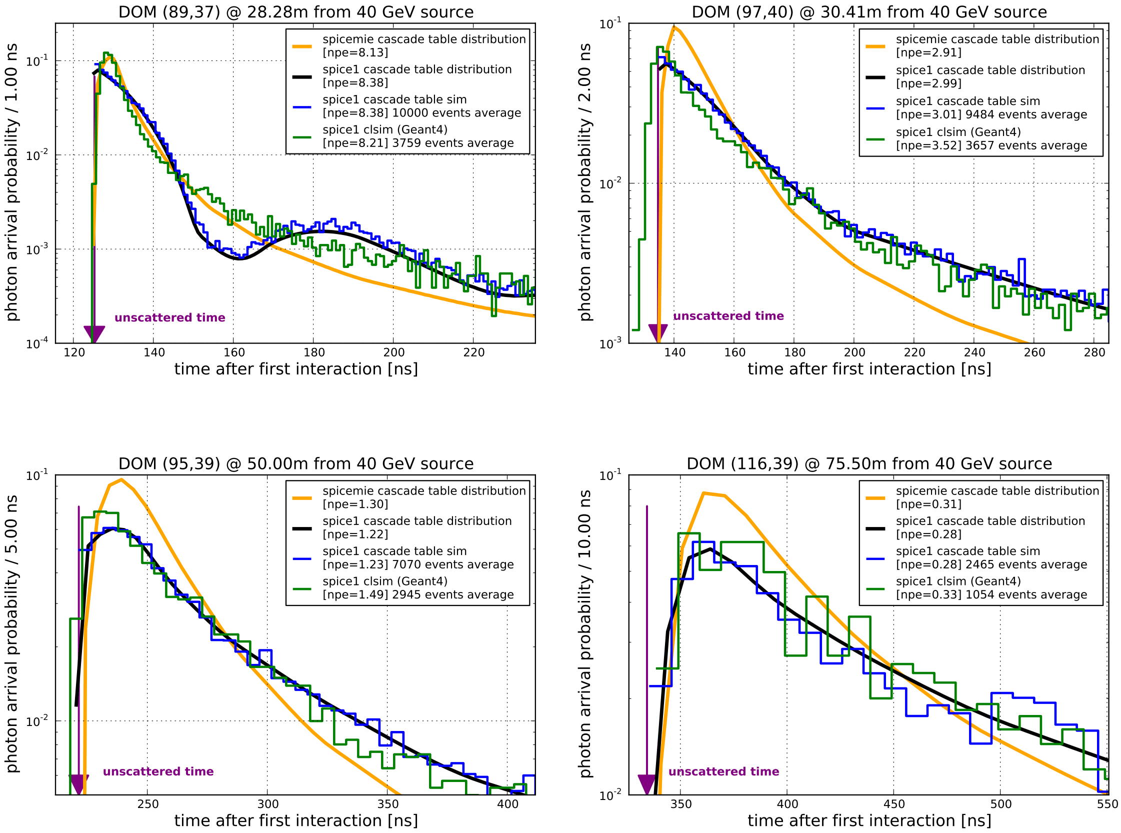

Figure 6.4.: Total amplitudes and differential photon arrival time distributions evaluated at increasing distances from an electromagnetic cascade of energy $E = 40\,\mathrm{GeV}$ placed at coordinates $x = 50\,\mathrm{m}$, $y = -52.2\,\mathrm{m}$, $z = -365\,\mathrm{m}$, close to the centre of the PINGU V15 detector volume. Times are taken with respect to the time of first interaction in the ice; PMT transit time spreads are not modelled. The total amplitudes include an assumed DOM efficiency reduction of 10 % due to cable shadow as well as relative quantum efficiencies, but not the PMT charge response distribution. In each figure, the black and orange lines represent the delay time quantiles as obtained from spline tables for the appropriate source-observer configurations, for two different ice models (Spice1 and SpiceMie). The result of a Spice1 simulation drawing amplitudes and times from the table can be compared to its parent distribution. In addition, results from a full photon propagation with `clsim` in Geant4 mode are displayed (green). At the smallest distance, the Spice1 spline representation exhibits additional structures above approximately 160 ns, which are found in none of the other distributions. While the total amplitudes can differ substantially, at larger distances, where the curves are less strongly peaked, full propagation and spline table distributions show somewhat better agreement.

7. Multinest Cascade Reconstruction Studies

Building on the Poissonian likelihood formulation of Eq. (6.3), the covariance matrix (6.7) provides an event-by-event a-posteriori estimate of the accuracy with which the true event parameters can be recovered in a maximum likelihood reconstruction, simultaneously taking into account possible correlations between different estimators. Being able to select high-quality events, i.e. those with small reconstruction residuals in the parameters of interest, can prove beneficial to physics analyses in PINGU, whose sensitivity to the neutrino mass hierarchy depends to a large extent on how well the energies and directions of atmospheric neutrinos traversing the Earth can be reconstructed [1, 22, 82].

Earlier detailed studies on the reconstruction of cascade-like events at energies above 100 GeV carried out in Ref. [74] revealed different approximations in the likelihood function that lead to the breakdown of the purely Poissonian description of the observed light yields, and demonstrated how their removal resulted in a consistent behaviour of resolution estimates obtained from directly sampling the likelihood space. Building on these results, this chapter will examine under which conditions resolutions of electron-neutrino induced CC events reconstructed in PINGU can be reliably predicted using Eq. (6.7).

7.1. Default reconstruction scenario

At the time this analysis was started, the PINGU baseline geometry was V15, which is an array of 40 strings with a horizontal spacing of 20 m, instrumented with 60 DOMs each, spaced 5 m apart in the vertical direction. For both older and more recent ice models, there were cascade spline tables that had been produced for reconstructing higher energy events in IceCube, while muon spline tables (for use with track segments of 15 m length, i.e. ~ 3.3 GeV minimum ionising muons) were available only for the somewhat outdated ice model Spice1. Official PINGU MC events were simulated with the more realistic SpiceMie ice model, with the drawback of having to use a less recent ice model for reconstructing tracks. Since in the following the likelihood itself is supposed to be analysed for its suitability in describing the observed light yields, it was decided to use the same ice model in simulation and reconstruction, Spice1. As is illustrated in Fig. 6.4 in an exemplary manner, even for a single ice model there are some discrepancies between directly simulated amplitudes and those which have been fit with splines.

For this data set, an isotropic flux of electron (anti-)neutrinos with energies between 1–80 GeV was generated according to an E^{-1} spectrum, and the detector response was simulated as laid out in chapter 4, with noise on all PINGU and IceCube DOMs and assuming a relative DOM efficiency decrease of 10 % due to cable shadow. The likelihood was maximised in eight dimensions[1] with Multinest, using cleaned pulses, a binning of 1 photon per bin, and a 2.19 ns Gaussian smearing of photon delay time distributions to account for the simulated PMT transit time spread.

In order to estimate how well the likelihood formulation describes the actually simulated detector response, it is instructive to compare the likelihoods of the reconstructed and true sets of parameter values, l_{reco} and l_{true}. First of all, assuming that the maximisation algorithm exactly locates the global likelihood maximum, since both likelihoods are evaluated using the same detector readout and identical reconstruction parameters, one would always expect $l_{\text{reco}} \geq l_{\text{true}}$, except for some cases where one or several of the MC truth parameter values lie outside of the allowed prior range for the corresponding reconstructed parameters. Then, Eq. (5.14) gives a probabilistic prescription of how far l_{reco} and l_{true} should lie apart when the likelihood formulation is an adequate description of the detector response. Before this comparison is possible, it is necessary to understand the way Multinest computes the likelihood of the true parameter values, based on the non-integer extension of Eq. (6.3).

7.1.1. Adaptation of the Multinest truth hypothesis to cascades

With the underlying event hypothesis being that of a ν_μ CC interaction (see Sec. 6.1), Multinest takes the secondary lepton as the true track with its corresponding energy and direction. The direction of the true cascade follows from conservation of momentum, and its energy corresponds to the visible fraction of the energy in the hadronic shower. Both track and cascade are constrained to originate from the true interaction vertex and time. Finally, the cascade and muon spline tables come into play by providing the bin-wise expectation values for all DOM-source combinations.

For ν_e CC events, this has as a consequence that the electron, whose length in the IceCube software framework is zero by definition, is not included in the event hypothesis, so that the latter will only be the hadronic shower. In the calculation of the likelihood, the muon spline tables will then also not enter. The ML hypothesis, however, will in general consist of both a track and a cascade, and will therefore make use of both types of look-up tables.

In order to render the two likelihoods more compatible, the definition of the true event parameters can be refined by adding the leptonic energy to the visible energy of the hadronic cascade and letting the direction of the new cascade coincide with that of the incident neutrino. In addition, the previous best fit likelihood is replaced by a cascade-only likelihood, which follows from adding the energy of the track reconstructed by Multinest to that of the reconstructed cascade and removing the track from the hypothesis altogether.[2]

The CC events selected for further analysis throughout the following studies had to satisfy a simple set of criteria: a log-likelihood exceeding the noise-only log-likelihood by 10, and containment of the reconstructed vertex within the PINGU fiducial volume, defined as a right cylinder centered at $x = 50\,\text{m}, y = -50\,\text{m}$ and ranging in depth from $-180\,\text{m}$ to

---

[1] In order to avoid a further increase in reconstruction time, it was decided to treat the cascade and track as collinear in this study.

[2] Or equivalently, setting the track length to zero. This functionality is already implemented in Multinest.

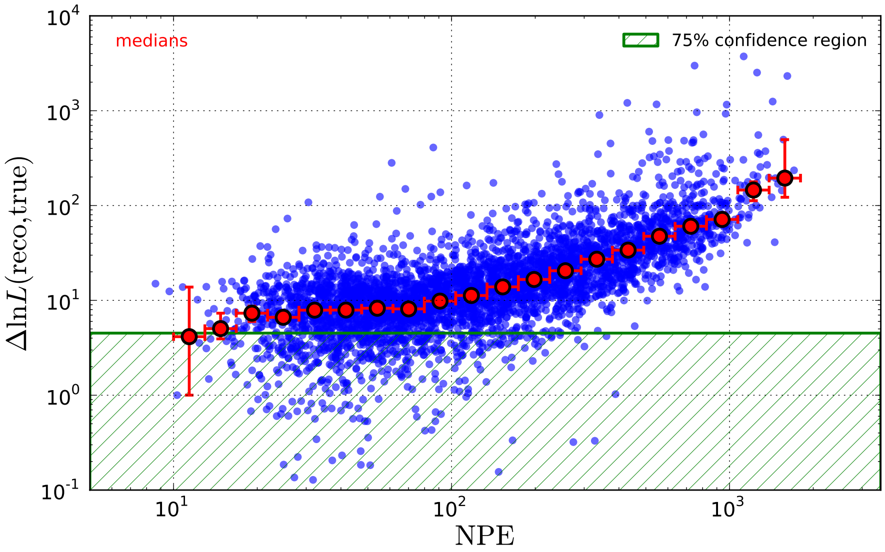

Figure 7.1.: Logarithm of the ratio between the Multinest likelihoods of the reconstructed and true event hypothesis parameters as a function of charge in the event (NPE) for the default simulation and reconstruction scenario, but with a refined truth likelihood and an "only-cascade" representation of the reconstructed hypothesis parameters. Red points show the median in each charge bin.

$-500\,\text{m}$. Further, to not retain events whose best fit parameters might be artificially far away from the actual likelihood maximum, it was required that the maximisation process had not been forced to terminate due to an upper time limit of $\sim 1.5\,\text{h}$ per event.

Fig. 7.1 shows the logarithm of the likelihood ratio $\Delta \ln L(\text{reco}, \text{true}) \equiv \ln\left(\frac{L_\text{reco}}{L_\text{true}}\right) = l_\text{reco} - l_\text{true}$ (from here on simply referred to as *log-likelihood ratio*) resulting from this refinement, as a function of total event charge (after noise cleaning) summed over all DOMs. For a reconstruction subject to statistical fluctuations only, the data should ideally fall into the hatched region defined by the condition $\Delta \ln L(\text{reco}, \text{true}) \leq 4.5$ with a probability of $\sim 75\,\%$, over the whole range of charges (neglecting effects related to small samples, which would modify the probability statement (5.14)). As one can see, in contrast to the ideal scenario, the likelihood ratio scales with charge over the whole range; the ML hypothesis becomes increasingly more likely than the hypothesis of the true parameter values. At charges around 1000 PE, the two likelihoods differ by a factor of $\sim e^{100}$.

It is conceivable that this behaviour will affect the potential of estimating the reconstruction residuals, since the region that defines the covariance matrix of ML estimators obviously does not contain the true values in the vast majority of events. Therefore, before the underlying incompatibility between simulated detector response and reconstruction is further looked into in Sec. 7.2, the performance of the hybrid residual estimator introduced in Sec. 6.3.2 is examined.

7.1.2. Performance of the hybrid residual estimator

For the default reconstruction described above, the left panel of Fig. 7.2 shows the median residuals in vertex depth z, $|z_{\text{reco}} - z_{\text{true}}|$ as well as the corresponding residual estimates, while the right panel shows analog quantities for the zenith angle reconstruction, as a function of charge. In addition, fits to both true and estimated residuals are included which allow for a deviation from a pure power-law in charge by a constant term.

The vertex depth residuals only scale with charge up to $\sim 100\,\text{PE}$, which corresponds to neutrino energies of ∼5–30 GeV (cf. Fig. A.1). Above, the vertex depth reconstruction ceases to profit from additional photons, and is clearly dominated by a systematic component of $\sim 0.8\,\text{m}$ in magnitude. The width of the likelihood maximum, however, would in principle allow for a much more accurate reconstruction. A similar behaviour is seen for the other vertex coordinates and vertex time.

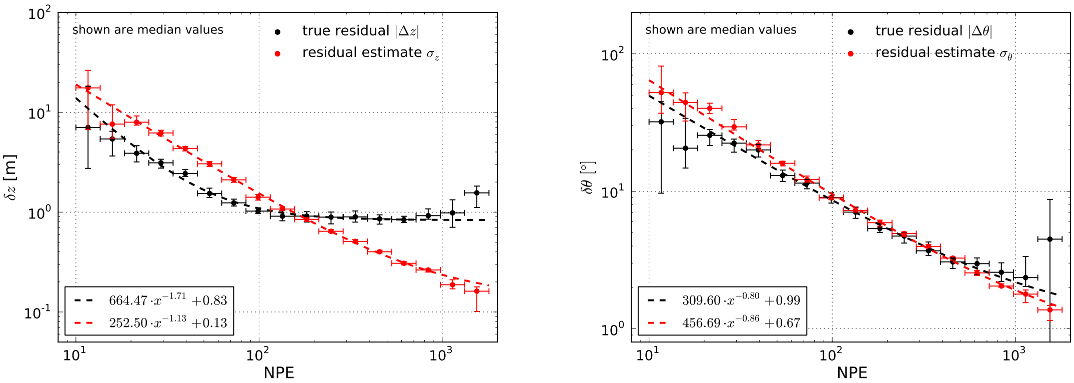

Figure 7.2.: Medians of true residuals (black points) and residual estimates (red points) for vertex depth z (left) and zenith angle θ (right) as a function of the observed event charge for ν_e CC events reconstructed within the PINGU fiducial volume by Multinest. Both quantities are parametrised as the sum of a power-law in charge and a constant term, thereby taking into account potential systematics impact.

In the case of the zenith angle reconstruction, there is better agreement between the residuals $|\Delta\theta|$ and the corresponding estimates σ_θ over the whole range of charges, even though the median $|\Delta\theta|$ begin to exceed the σ_θ at charges around 500 PE.

In order to examine the relation between the true residuals and their estimates in somewhat more detail, it is useful to evaluate the *pull* distributions [83]

$$P_a = \frac{a_{\text{reco}} - a_{\text{true}}}{\sigma_a} \,, \tag{7.1}$$

where a is one of the reconstructed parameters. If the ML estimates a_{reco} are indeed Gaussian centered about the respective true values a_{true} (see Sec. 5.2), then the pull distributions should also be Gaussian with mean $\mu = 0$. In addition, for suitable estimates σ_a, the pull distributions are expected to have standard deviation $\sigma = 1$, i.e. be standard Gaussian.

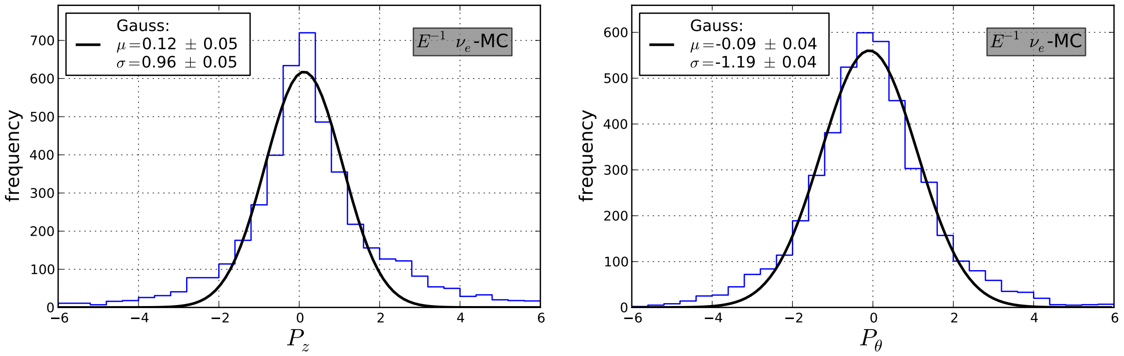

Figure 7.3.: Gaussian fits to the pull distributions in vertex depth, P_z, and neutrino zenith angle, P_θ, Eq. (7.1), for ν_e CC E^{-1} MC events reconstructed by Multinest. See text for details.

One expects this relation to only hold for a large number of observations, though, when the ML estimators acquire the demanded properties. Extended tails in the pull distributions, in contrast, can be manifestations of the non-asymptotic nature of the underlying problem [84].

Gaussian fits to the pull distributions P_z and P_θ are shown in Fig. 7.3. One can see tails in the left and right wings of the distributions which are not very well described by a Gaussian[3], but restricting the fits to the central-most region would lead to underestimated widths. The reconstruction is somewhat biased towards larger z and smaller θ, and the residual estimates σ_θ on average underestimate the true residual.

Since it has been found that residuals and estimates scale differently with photon yield, a more complete picture is obtained by examining the parameters of the pull distributions dependent on charge. Means and widths resulting from Gaussian fits to the P_a in different charge bins, again for vertex depth and zenith angle, are shown in the left and right panel of Fig. 7.4, respectively.

At small charges, the width of P_z is around 0.6–0.9 , indicating that the actual residual is overestimated. Above 100 PE, owing to the levelling out of the residuals, the widths become increasingly large, reaching $\sigma \approx 5$ for the highest charge bin. From Fig. A.2 in the appendix one can see that the distributions can be considered Gaussian above around 100 PE.

The overall trend of the width of P_θ is similar. While the estimates are too pessimistic at charges between 10–30 PE, their predictions are again significantly smaller than the true residuals for events in which large numbers of photons are observed. Note, however, that the scale of the effect is reduced in comparison to the vertex position reconstruction. Furthermore, apart from in the lowest-charge bin, reconstruction has the tendency to underestimate the neutrino zenith angle.

---

[3] For a discussion of models that are more suitable but harder to interpret see Ref. [83].

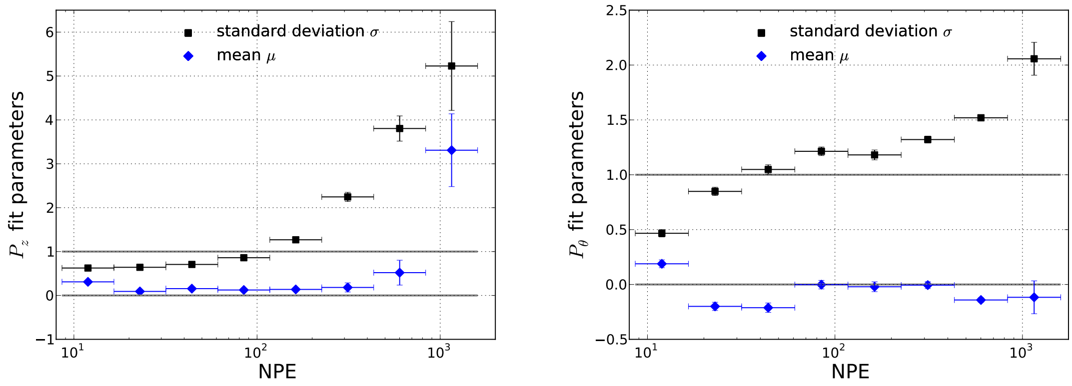

Figure 7.4.: Charge dependence of the pull distribution parameters for vertex depth z (left) and zenith angle θ (right). Shown are the means μ (in blue) and widths σ (in black) obtained from Gaussian fits. The gray horizontal lines correspond to the ideal values for an unbiased reconstruction whose residuals are on average perfectly predicted.

7.2. Single cascade reconstruction

Even with the refining modifications to the true and estimated event hypotheses and the corresponding likelihood evaluation in place, reconstruction still made use of both cascade and muon spline tables, while the true likelihood was evaluated based on the cascade spline tables only. While this asymmetry is intrinsic to cascade reconstruction with Multinest in hybrid mode, it makes it hard to find the causes of possible disagreement between simulation and reconstruction. Consequently, in the following Multinest was forced to assume a single cascade, thereby disabling the use of track spline tables in reconstruction. Simultaneously, the dimensionality of the likelihood maximisation procedure was reduced from eight to seven. It was found that the resulting log-likelihood ratios behaved almost identically to those shown in Fig. 7.1, with absolute values and scaling with charge of residuals and corresponding estimates (this time based on a single cascade hypothesis, cf. Sec. 6.3.1) very similar to Fig. 7.2. It is thus essential to discuss some persisting, more or less subtle inconsistencies between simulation and reconstruction, that might render the observed data less compatible with the true event hypothesis in this particular reconstruction scenario.

7.2.1. Deviations between simulated and modelled detector response

Cherenkov light yields and shower extension Most obviously, the amplitudes and photon delay times resulting from the direct propagation of secondary particles and photons with `clsim` in Geant4 mode do not necessarily agree with those stored in the `Photonics`-based spline tables. That this is indeed the case has already been demonstrated for the light yields at different distances from an exemplary cascade, Fig. 6.4. The latter also shows that, for a given source-observer configuration, the probability that a photon stemming from the full `clsim` particle propagation can reach a DOM prior to the geometric time based on the point-source assumption can be at the percent level.

Fig. 7.5 seems to indicate that the finite spatial extent of the simulated cascade is a limiting factor to vertex reconstruction accuracy. Depending on the visible energy of the true cascade, $E_{\text{vis,true}}$, as defined in Sec. 7.1.1 above, it depicts the median displacement of the reconstructed vertex from the true neutrino interaction point and from the approximate position of maximum energy deposition, respectively, where the latter is obtained using the longitudinal shower profile parametrisation (3.11) and fit parameters given in Ref. [27]. The cascade is clearly reconstructed significantly closer to the estimated position of maximum Cherenkov light emission than to the true vertex. Also, increasing energy—and thereby shower depth $z_{\max}$—results in a larger difference between the medians of the two distributions.

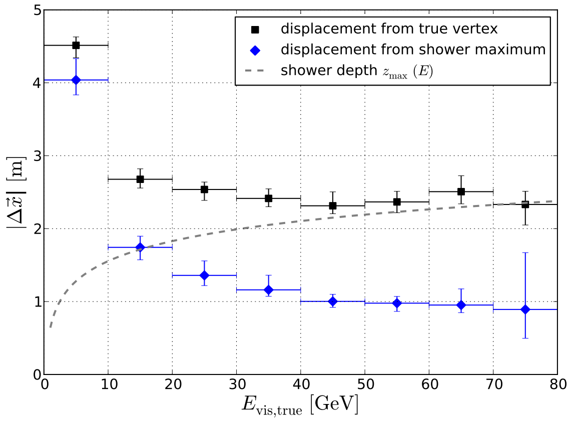

Figure 7.5.: Median distances between the (point-like) reconstructed cascade and true ν_e interaction vertex (black squares) or estimated position of maximum shower development (blue diamonds), respectively, as a function of visible energy. For reference, the gray dashed line shows how the depth of maximum energy deposition of an electromagnetic cascade develops with energy, using fit parameters from Ref. [27].

Another potentially important contribution arises from the parametrised relative Cherenkov light yield suppression of a hadronic cascade of given energy. In Multinest, it is assumed that the true electromagnetic equivalent energy of the hadronic cascade originating from the vertex of the ν_e CC interaction is actually given by this parametrisation of the average cascade development. Event-to-event fluctuations around the average are considerable, however, exceeding 20–30 % in the GeV energy regime [33]. This in turn is expected to result in less agreement between the observed amplitudes and the true parameters of the event.

Noise Furthermore, while the likelihood formulation assumes a (constant rate) contribution due to noise, the pulse cleaning (ideally) removes all noise hits. When noise hits are removed, a fraction of the physical charge might be interpreted as being due to noise, and as a consequence the true event parameter values might be suppressed in favour of others which yield a smaller amplitude. Even if noise hits were retained, the constant noise rate hypothesis would deviate from the correlated noise that is simulated in addition to the thermal component.

Further approximations A number of additional approximations present in the likelihood function include the attempts to account for fluctuations arising from the PMT charge response distribution by simply scaling the expected “raw” light yields by the mean of the distribution (see discussion in Sec. 4.3), and to account for fluctuations arising from the (only nearly Gaussian) PMT transit time distribution by convolving the mean photon arrival time distributions by a Gaussian of similar width. At the base of these approximations lies the assumption that PMT pulse time and charge distortions induced by waveform calibration and pulse extraction are negligible. Finally, saturated DOMs can also lead to a bias away from the true parameter values when only single time intervals during which saturation occurs are excluded, and not the whole affected DOM. For an extensive discussion of these and further systematics see Ref. [74].

7.2.2. No photon arrival time binning

A reconstruction carried out on all detected pulses (including noise) did not show any perceptible change in the log-likelihood ratio behaviour encountered above, except for an expected shift to higher overall event charges. In order to disentangle effects that might stem from an erroneous description of photon delay times and those originating from wrong assumptions about the amplitudes, temporal binning of pulses was disabled. In this case, the whole readout range acts as one large bin for each DOM, and only the total amplitudes expected and detected within this time window are used in evaluating the likelihood. For this reason, saturated DOMs are automatically not considered in reconstruction. The log-likelihood ratios depicted in Fig. 7.6 exhibit a much less rapid increase with NPE than in the case when the PMT pulses are binned in time, but the significant undercoverage of the region with ideally 75 % probability content persists: it contains only 14 % of all events.

In the meantime, it had been reported that the spline tables used in PINGU, but created for events with energies of TeV and above, suffered from somewhat poor fits to the raw photon amplitudes (which they had been originally generated from) at small distances of the order of several metres from the source, both for muon and cascade tables. Hence, it might turn out benefitial to exclude charge detected by the DOMs closest to the true vertex (e.g. within some radius R) in reconstruction, in the sense of obtaining an improved modelling of the simulated detector response. Here, one should make sure that these DOMs are not considered unhit in the likelihood formulation, e.g. by adding them to the list of broken DOMs. Indeed, the charge dependence of the log-likelihood ratios in Fig. 7.7, which is the result of excluding all DOMs within a sphere of radius $R = 30\,\mathrm{m}$ centered at the true interaction vertex, all but disappears, except for perhaps the last one or two charge bins.

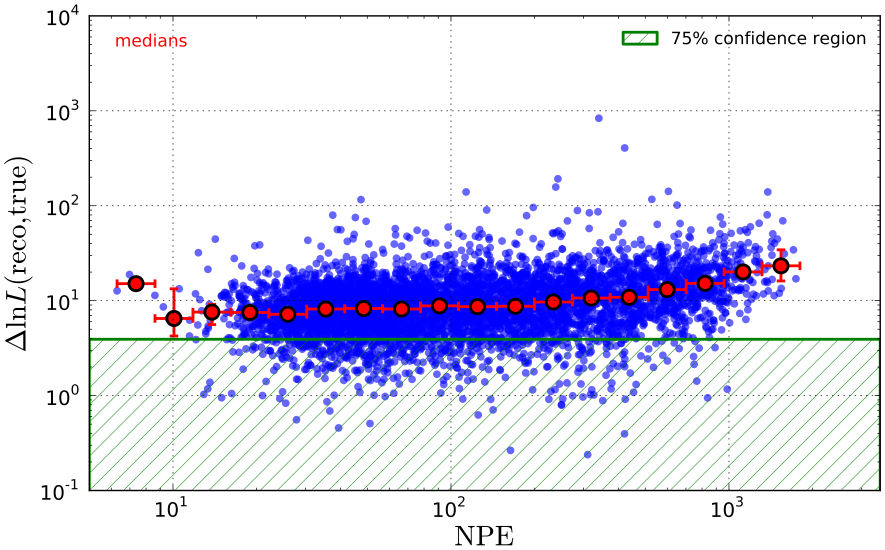

Figure 7.6.: Log-likelihood ratio for a realistic cascade simulation and an only-cascade reconstruction which does not make use of photon arrival time information.

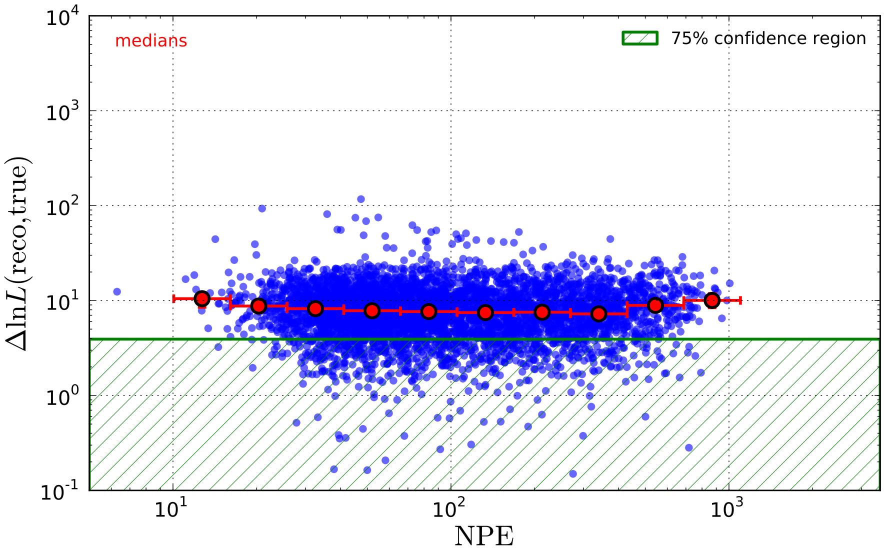

Figure 7.7.: Same as Fig. 7.6 , but without using information from DOMs that are located within a sphere of 30 m radius centered at the true interaction vertex.

7.3. Idealised simulation using spline tables

A simple approach that should mitigate differences between the true and expected Cherenkov light yields is the generation of photons based on the same spline tables that are used in reconstruction. For each light producing secondary in the `I3MCTree`, the number of photons at any given DOM location is drawn from a Poisson distribution whose mean is extracted from the tabulated light yields, and residual times, Eq. (6.5), are obtained based on the delay time distribution for the specific source-observer configuration. If the secondary in question is a hadron, it is first assigned a reduced effective energy, simultaneously taking into account the energy-dependent relative fluctuations around the average light output. In reconstructing real events, this would correspond to perfect knowledge of ice properties and the Cherenkov light output of a hadronic cascade being independent of the type of hadron that initiated it.

The simulated events were used to perform different single-cascade reconstructions with Multinest, whose corresponding median log-likelihood ratios for each charge bin are depicted in Fig. 7.8. Here, one can compare the results of two reconstructions using timing information, one including a 2.19 ns convolution of the expected photon arrival times and one without, fully excluding DOMs from the likelihood evaluation if parts of their waveforms are marked as invalid. In addition, both of these are contrasted with an amplitude-only reconstruction. Since the individual data points are not shown, the fraction of events that is covered by the confidence region defined by the quantile of order 0.75 of the χ^2 distribution with 7 and 6 d.o.f., respectively, is given in the upper right-hand corner.

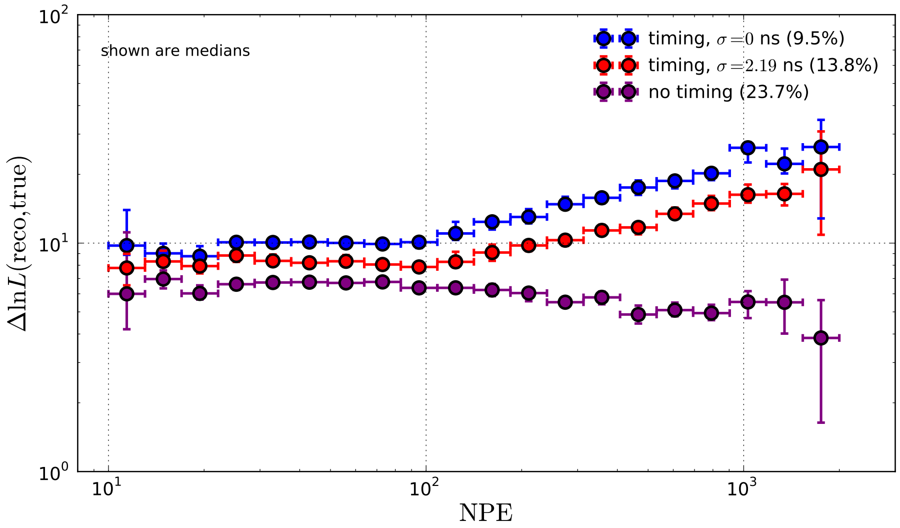

Figure 7.8.: Median log-likelihood ratios for three different reconstructions of ν_e CC events simulated using tabulated light yields. In the case of the purple sample, the likelihoods were evaluated without incorporating pulse times. The red and blue samples, on the other hand, are the result of binning PMT pulses in time, with the difference that in the case of the former photon delay time distributions were additionally convolved with a Gaussian of $\sigma = 2.19$ ns width. See text.

The overall behaviour of the two data sets that use timing information is similar: the log-likelihood ratio is approximately constant up to charges around 100 PE and then starts to rise. As expected, a somewhat improved description of the detector response is obtained by modelling PMT transit time spread. Both the magnitude of the log-likelihood ratio as well as its charge dependence resemble those seen in the amplitude-only reconstruction for the more realistic simulation, Fig. 7.6.

Not making use of timing information leads to a log-likelihood ratio that is also quite flat at charges below 100 PE—albeit with smaller values of $\Delta \ln L$ since the likelihood in this case is independent of the time of the source—but which exhibits a small negative slope for larger amplitudes, implying increasing agreement between true and reconstructed parameters where the relative fluctuations expected from a purely Poissonian process become smaller than $\sim 10\,\%$. Nevertheless, the fraction of events that can be found in the region of ideally 75 % coverage still only amounts to $\sim 24\,\%$ when averaged over all charges.

Fig. 7.9 allows one to directly compare the true residuals and corresponding estimates again for vertex depth and neutrino zenith angle for those events reconstructed under the use of timing (with $\sigma = 2.19\,\mathrm{ns}$, referred to as "timing-sample" in the following) to those obtained for the amplitude-only reconstruction. Here, the covariance matrices were computed according to the prescription given in Sec. 6.3.1, i.e. under the assumption of a single-cascade hypothesis. In general, the estimates succeed relatively well in predicting the reconstruction residuals for both samples. In the case of the timing-sample, one can still observe a levelling-off of the z-residuals at the high end of the charge range, but reconstruction is less strongly affected by systematics than in the more realistic direct simulation scenario, Fig. 7.2. The estimates of the amplitude-only residuals perform best at highest charges.

In addition, it is important to note that it is not only the degree of consistency between simulation and reconstruction that governs how accurately the true parameters can be recovered, but also the amount of information that is used. The cascade's forward-peaked photon emission in the direction of the Cherenkov angle depicted in Fig. 3.3, for example, becomes less and less pronounced as time evolves due to photon diffusion, so that the hit pattern exhibits increasing spherical symmetry; see e.g. Ref. [85] for a discussion. Despite the scaling of the log-likelihood ratio with charge, vertex depth and zenith angle resolutions are better when the likelihood incorporates photon arrival times, at least for the charges encountered here. At 500 PE, adding timing information improves the accuracy with which the true vertex depth z is reconstructed from around 1.1 m to slightly below 0.3 m, and neutrino zenith angle resolution from 7° to 3°. Interestingly, for the timing-sample, the functional dependency of the median zenith angle residual on charge is very similar to that of the hybrid Multinest reconstruction in the default scenario discussed in Sec. 7.1.2, where the secondaries constituting the hadronic and electromagnetic cascades were fully propagated through the ice. This could indicate that directional reconstruction profits from the spatial extent of the source in the full simulation.

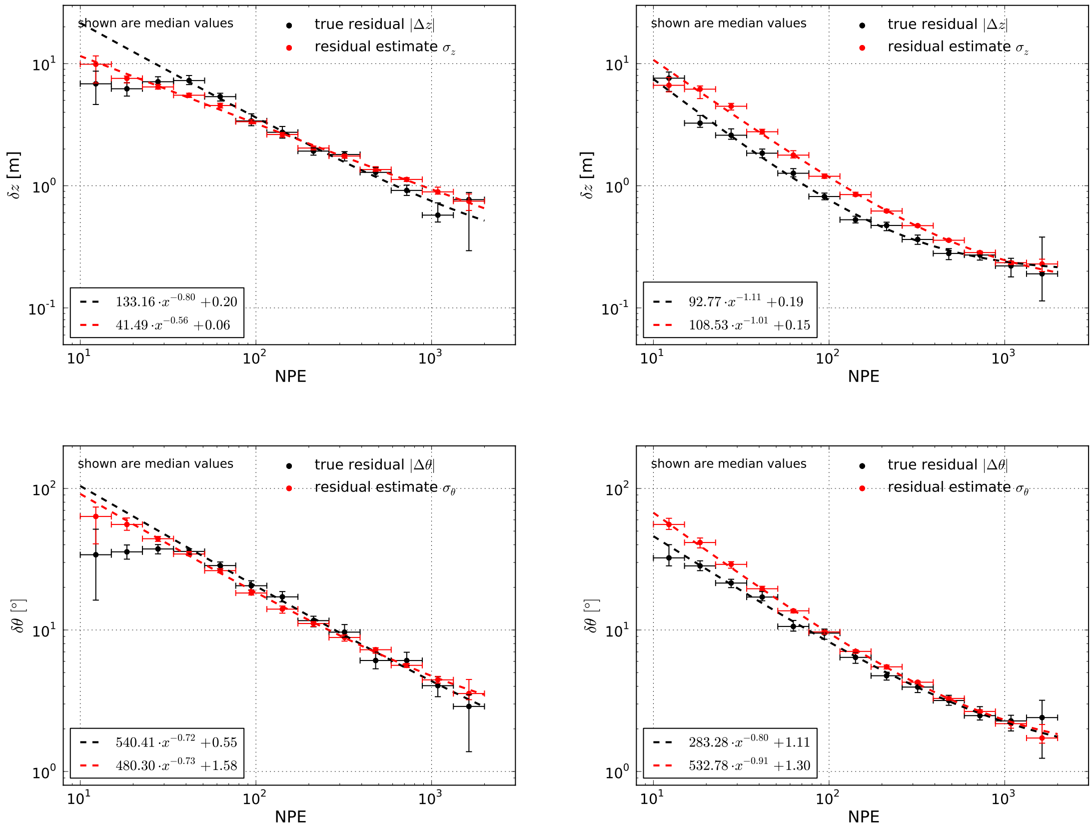

Figure 7.9.: Same as Fig. 7.2, but for a simplified simulation using cascade spline tables and a reconstruction assuming a single-cascade hypothesis. The two left panels show residuals and estimates for vertex depth and neutrino zenith angle when photon arrival times are ignored, while the same quantities in the two panels on the right-hand side were obtained by taking times into account and convolving the delay time distributions with a Gaussian of width $\sigma = 2.19\,\text{ns}$.

7.3.1. Removal of PMT effects and uncertainties in hadronic light output

In order to be able to judge whether discrepancies between predictions and true residuals seen above are due to intrinsic limitations related to the non-asymptotic nature of the ML estimation problem, or whether they arise from remaining implicit, unrealistic assumptions about the detector response in the likelihood formulation, it is necessary to further break down the whole process of simulation and reconstruction.

On the simulation side, this was achieved by removing the neutrino scattering process and instead directly placing electromagnetic cascades in the ice, which have maximal Cherenkov light output. Photoelectrons generated based on spline table predictions, each with an assigned charge of one and a given arrival time, were then directly used as an input to reconstruction.[4] Also, no photoelectrons resulting from noise were simulated. The likelihood formulation was then modified to expect exactly 1 PE per photon instead of the 0.86 PE that correspond to the mean of the SPE PMT charge response distribution, and the assumed noise rate[5] was scaled down to 1 Hz. Given this scenario, with all other relative efficiency correction factors agreeing (see Sec. 6.2.2), the likelihood (6.3) should represent a suitable model of the physics of the underlying event.

The log-likelihood ratios resulting from an amplitude-only Multinest reconstruction based on the cascade hypothesis can be examined in Fig. 7.10. Averaged over the whole charge range, 76.3 % of the events for which $l_{\text{reco}} > l_{\text{true}}$ fall into the hatched region, indicating good agreement with the ideal coverage of 75 % for a purely Gaussian likelihood. Due to the finite accuracy of the maximisation algorithm, it is not unexpected that a confidence region of nominal probability content β in the asymptotic limit, defined via Eq. (5.14), exhibits an overcoverage $\beta' > \beta$.

Since the error estimates should yield one-sigma confidence intervals only in the asymptotic limit, their accuracy necessarily depends on the degree to which this limit is reached indeed. As the likelihood is known to have the beneficial property of being "strictly concave, with a single maximum" [52] in energy, a small study was devised in which only the visible energy of the cascade was reconstructed, while the other parameters were kept fixed at their true values. The variance on the energy estimator $\hat{E}$—when only total amplitudes and no arrival times are observed—then follows from Eq. (6.7) as

$$\mathrm{Var}(\hat{E}) = \left(\sum_D \frac{1}{\mu_D} \left.\frac{\partial \mu_D}{\partial E}\right|^2_{E=\hat{E}}\right)^{-1} = \left(\sum_D \frac{\Lambda_D^2}{\mu_D}\right)^{-1} , \tag{7.2}$$

where in the second step the linear dependence between Cherenkov amplitude and cascade energy, Eq. (6.2), was exploited. If the mean amplitude μ_D has no contribution from noise, we can use relation (6.2) once again to cancel a factor Λ_D and finally write

$$\mathrm{Var}(\hat{E}) = \left(\sum_D \frac{\Lambda_D}{\hat{E}}\right)^{-1} = \frac{\hat{E}}{\sum_D \Lambda_D} . \tag{7.3}$$

---

[4] More precisely, each photoelectron (`I3MCPE`) first was transformed into an `I3RecoPulse`, with identical time, a charge of one and a given width of 3.3 ns, corresponding to the typical SPE pulse width.

[5] A finite value is required for numerical stability [75].

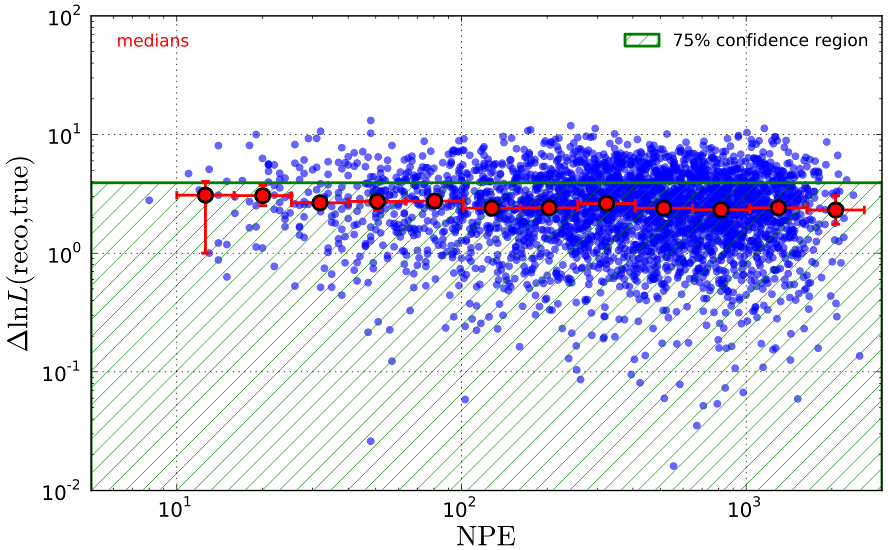

Figure 7.10.: Log-likelihood ratio for an amplitude-only reconstruction of point-like electromagnetic cascades, with all non-Poissonian fluctuations removed.

Note that the denominator is determined by the depth and zenith angle of the source as well as the deplacement vectors that connect it to the various receivers D, but is independent of energy (cf. Sec. 6.2.2).

Fig. 7.11 shows by how much the likelihood has reduced with respect to its maximum value $l_{\hat{E}}$ at the bounds of the confidence intervals $\hat{E} \pm \sigma_{\hat{E}}$, $\sigma_{\hat{E}} = \mathrm{Var}(\hat{E})^{1/2}$, depending on the total charge[6] detected in the event; the gray line at 0.5 represents the expectation in the asymptotic limit. The relative deviation decreases from around 5 % for the smallest charges to 1 % at 50 PE, and becomes negligible for amplitudes exceeding 1000 photoelectrons.

Complementary information about the relative asymmetry between the values of the likelihood at the one-sigma bounds is provided in Fig. 7.12: a positive ratio implies that the likelihood falls off more rapidly toward small energies, which is the case for all events with charges below 300 PE, and which can also be observed in the low-energy cascade example of Fig. 6.3. Above, the likelihood tends toward perfect symmetry.

Results of full reconstructions in all six dimensions and the evaluation of the diagonal elements of the covariance matrices are shown in Fig. 7.13, for zenith angle and visible energy. The charge-independent summands in the fit were set to zero to avoid non-physical negative predictions for the $|\Delta\theta|$ and σ_θ in the asymptotic limit, which resulted from the rather flat behaviour at very small charges. The fits to the fractional energy residuals and error estimates, however, were not influenced by this constraint. Both median zenith angle residuals, which decrease from around 26° at 50 PE to 2° at 3000 PE, as well as fractional

---

[6] Here, due to the specific nature of the simulated experiment, “charge” and “number of photons” can be used interchangeably.

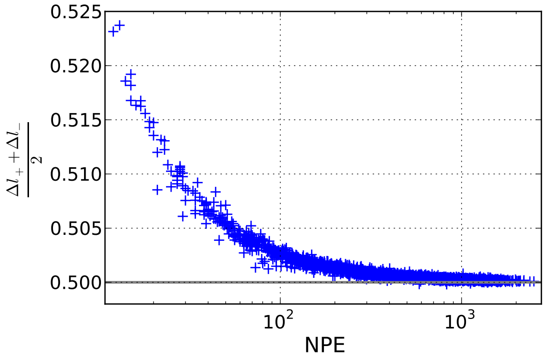

Figure 7.11.: Mean difference between the likelihood maximum $l_{\hat{E}}$ and the values at the bounds of the symmetric interval $\hat{E} \pm \sigma_{\hat{E}}$, dependent on the number of photons that are detected by all DOMs in the event.

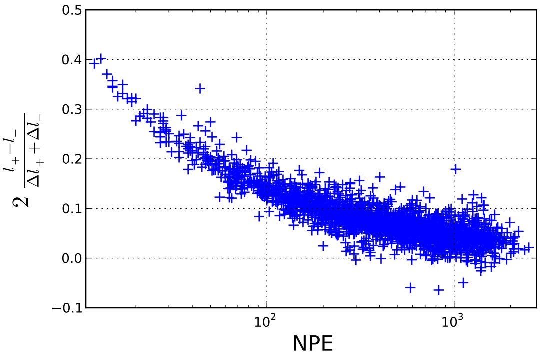

Figure 7.12.: Relative difference (asymmetry) of the values of the likelihood at the bounds $\hat{E} \pm \sigma_{\hat{E}}$ of the symmetric interval about the cascade energy estimator $\hat{E}$.

energy residuals, which become as small as 1 % for large numbers of detected photons, are clearly dominated by statistics.

Means and standard deviations of Gaussian fits to the corresponding pull distributions are depicted in Fig. 7.14. No clear picture emerges for the zenith angle pulls: the means are consistent with zero, and the standard deviations seem to approach the desired value of one as photon statistics accumulate. The energy estimator, however, seems to have a negative bias. In the limit in which $\hat{E}$ is uncorrelated with other estimators, Eq. (7.3) then shows that the variance estimate is small, which leads to the expanded pull scales at small charges. As the bias in energy diminishes with an increasing number of observations, the pull distribution approaches a standard normal.

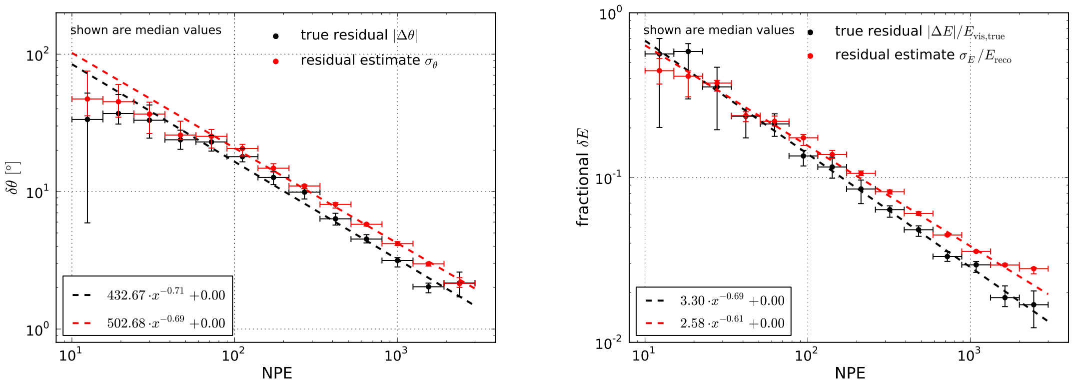

Figure 7.13.: Same as Fig. 7.2, but for an idealised amplitude-only reconstruction of point-like cascades with all non-Poissonian fluctuations removed. The left panel shows residuals and estimates for cascade zenith angle, while the right shows the same for fractional cascade energy resolution. Refer to text for details.

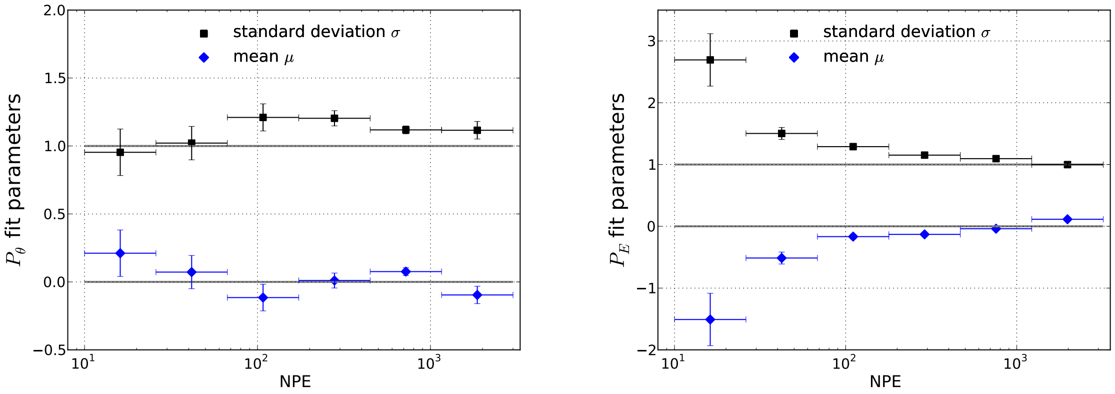

Figure 7.14.: Charge dependence of the pull distribution parameters for cascade zenith angle (left) and visible energy (right) for the idealised amplitude-only reconstruction of point-like cascades. Shown are the means μ (in blue) and widths σ (in black) obtained from Gaussian fits. See text for details.

7.4. Photon timing schemes

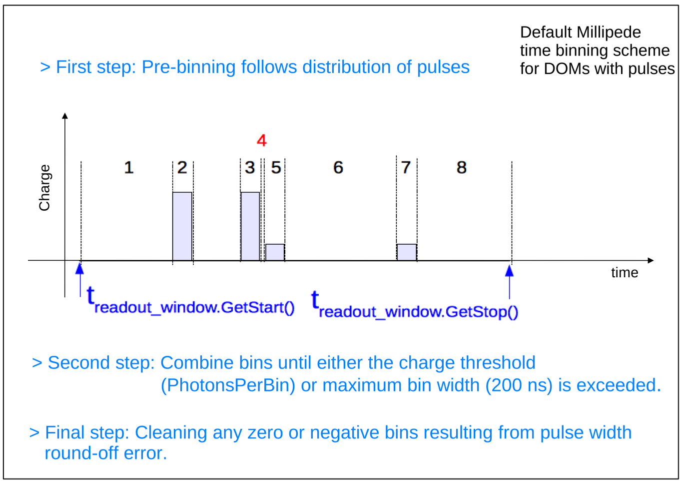

Figure 7.15.: Essentials of the default `Millipede` time binning scheme for a DOM with PMT pulses used in this work. A global readout window determines the earliest and latest pulse times that are considered, and the threshold is the desired charge in each bin. The numbers denote the different time bins. Figure taken from Ref. [86].

As timing was added back in and the resulting behaviour of the log-likelihood ratio was examined, it was surprising to see that it showed a strong scaling with charge again (see Fig. A.3). If this had been due to an offset in the expected amplitudes, a similar effect should have also showed up in Fig. 7.12, i.e. it must have had something to do with the binning process itself.

In Fig. 7.15 the concept behind the time binning of PMT pulses in `Millipede`, as used for these studies up until now, is outlined. A global readout window defines the earliest and latest times that are taken into account. In a first step, the pulses detected by any given DOM are pre-binned according to their start times and widths. Depending on user options, i.e. the desired amount of charge per bin, bins are then combined until either the charge content of the bin is larger than the threshold or the duration of the bin exceeds a maximum width of 200 ns.

Furthermore, each DOM can be assigned one or several time periods which should be excluded, for example because the waveforms exhibit saturation. Bins that fall into one of these intervals are declared as invalid and do not contribute to the likelihood. This approach was, as mentioned in Sec. 7.2.1, not taken throughout this work, though, and instead affected DOMs were fully excluded from the likelihood evaluation. As in the case of an amplitude-only reconstruction, for DOMs without any detected charge the readout window is not subdivided in time.

While this binning scheme provides a good trade-off between precision and computational demands, it can lead to bin widths for single DOMs ranging from the sub-ns to the µs regime, and an irregular likelihood-behaviour has been associated with it [86]. Also, it is questionable whether the assumption that bin entries are distributed Poissonian around the

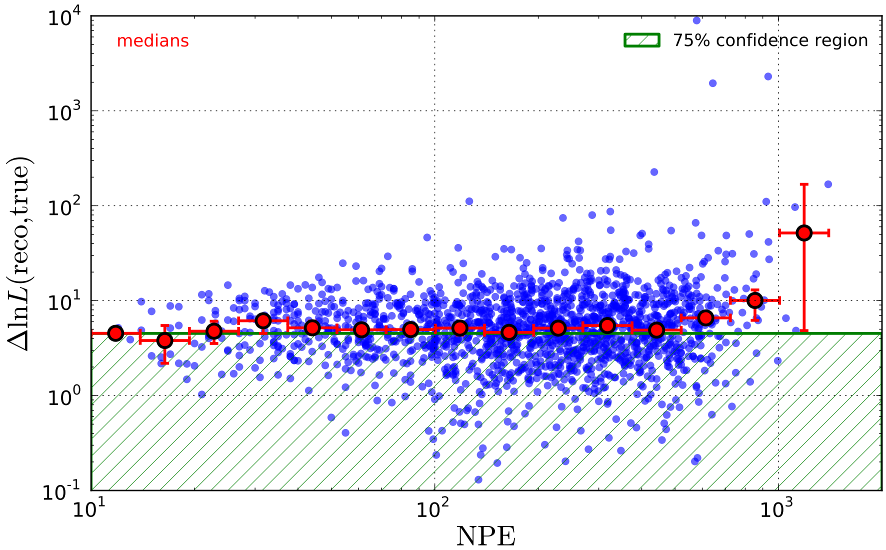

Figure 7.16.: Same as Fig. 7.10, but using a photon timing scheme that makes use of bins of width 5 ns.

expected mean still holds, as the longest bins will have no detected charge by construction.

In order to test the effect that time binning has on the log-likelihood ratio between reconstructed and true parameters, reconstructions based on an implementation of `Millipede` that makes use of equally spaced bins [87] were performed. For each DOM, an individual readout range starting 1 µs prior to the occurrence of the first and ending 1 µs after the occurrence of the last photon was defined. The resulting readout ranges of 2 µs length were then subdivided into intervals of 40 ns, 20 ns or 5 ns width, corresponding to 50, 100 and 400 bins, respectively.

Fig. 7.16 shows the result of the reconstruction with bins of 5 ns width. The median log-likelihood ratio is very flat; 45 % of the events lie within the hatched region, which is significantly less than the ideal value of 75 %, but shows clear improvement over the results obtained with the default binning scheme. In addition, while the ratio stayed constant over the whole charge range, it was observed that it shifted upward with increasing bin size, i.e. the values of the parameters constituting the true event hypothesis became less compatible with the observed light yields and arrival times.

A similar behaviour of the log-likelihood ratios seemed to emerge also for a more realistic scenario, with electronics effects included, but due to the large computational expense that accompanied this binning scheme, this could not be followed up in more detail.

8. Application to Neutrino Mass Hierarchy Sensitivity

Estimates of the PINGU sensitivity to the NMH have found that the major contribution arises from the cascade channel, which is due to larger statistics and a reduced systematics impact compared to the track channel [23]. The importance of a precise reconstruction of the neutrino energy and zenith angle in the determination of the NMH with large atmospheric neutrino detectors has already been motivated in Sec. 2.4. In the previous chapter, the accuracy of a likelihood-based, semi-analytic approach to estimate the intrinsic resolution of an event has been shown to depend on how well the actual detector response is modelled. The most realistic scenario of Sec. 7.1/7.2 hinted at the existence of systematics that dominated over Poissonian fluctuations in amplitude at high charges, which resulted in resolution estimates that were too optimistic.

The following will present an application of the resolution estimator to the measurement of the NMH with PINGU geometry V36 (see Sec. 3.5.1), where only the electron neutrino channel ($\nu_e + \bar{\nu}_e$ CC) is considered. First, the analysis framework and the Monte Carlo data set used are briefly introduced, before simple cuts on the resolution estimates resulting in event samples of varying quality and statistics are illustrated. Finally, the uncertainties of different parameters and their impact on the NMH significances are discussed.

8.1. Parametric PINGU Analysis

A set of independent, complementary statistical analyses, with focus on either accuracy or computational time, aim at evaluating PINGU's NMH sensitivity. Among these, the *Parametric PINGU Analysis* (PaPA) [23] is a particularly fast approach that can take into account a wide range of systematic effects with the potential to impact the sensitivity of PINGU. It is based on the *Fisher information matrix*

$$F_{kl} = \sum_i \frac{1}{\sigma_{n_i}^2} \frac{\partial n_i}{\partial p_k} \frac{\partial n_i}{\partial p_l} \bigg|_{\text{fid. model}} , \tag{8.1}$$

computed from the (numerical) partial derivatives of the number of events in each $(E, \cos\theta)$ bin i with respect to the set of parameters p_i under study. The derivatives in Eq. (8.1) are evaluated at the values of p_i constituting the model which represents the best existing knowledge and is used as an input for the simulation of the n_i. Inversion of the Fisher information matrix (8.1) yields the full covariance matrix of the experiment, from which one can obtain the full uncertainties as well as correlations between different parameters.

The physics parameters include three mixing angles, two mass-squared differences, the

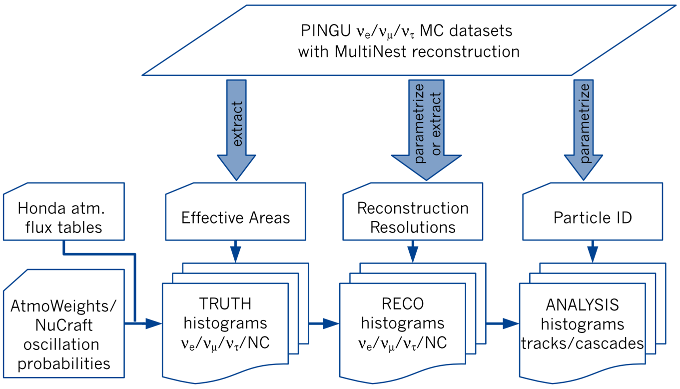

Figure 8.1.: Parametric PINGU Analysis (PaPA) simulation flow chart [23]. MC generated events are used to parametrise the effective detection areas, reconstruction resolutions and flavour identification. See text for details.

CP-violating phase and a hierarchy parameter h. For treatment with the Fisher matrix, the latter is defined in a continuous way: $n_i(h) = hn_{i,\mathrm{NH}} + (1-h)n_{i,\mathrm{IH}}$, $0 \leq h \leq 1$, where $h = 1$ corresponds to the NH and $h = 0$ to the IH. As the result of a study showing negligible impact of the solar mixing parameters $\Delta m^2_{\mathrm{sol}}$, θ_{sol} and the CP-violating phase δ_{CP}, these parameters are kept fixed at their respective fiducial values.

Fig. 8.1 shows a scheme of the various steps that constitute the general analysis procedure. Each step results in a histogram in the $(E, \cos\theta)$ plane which takes into account an increasing number of experimental signatures. First of all, the histograms containing all the relevant oscillation probabilities of atmospheric $\nu_{e,\mu}/\bar{\nu}_{e,\mu}$ as a function of energy and zenith angle are created, under the assumption of the Earth model introduced in Sec. 2.3. Models describing the atmospheric neutrino fluxes are then employed to transform the oscillation probabilities into fluxes for $\nu_{e,\mu,\tau}/\bar{\nu}_{e,\mu,\tau}$ arriving at the detector, which in turn are multiplied by the corresponding effective areas to obtain event rates. The last step that relates to the studies carried out in this work is the application of reconstruction resolutions to the histograms of true event numbers by smearing the bin counts accordingly. Especially, no background is assumed from either ν_τ CC events or from ν_μ events when the muon cannot be identified as such, nor from NC events in general. Conversely, no electron neutrinos are assumed to be misidentified.

8.2. Data set

The data set from which the effective areas and resolutions for $\nu_e/\bar{\nu}_e$ CC events are obtained was simulated with PINGU geometry V36 (run 356). An isotropic flux of neutrinos was generated with GENIE according to an E^{-1} spectrum, in the energy range 1–80 GeV. Secondary interaction particles and photons were propagated in SpiceMie ice via direct simulation. The eight-dimensional Multinest reconstruction was performed, fitting for both a cascade and a collinear track, using SpiceMie spline tables for "point-like" muons—optimised for low energy reconstructions—and cascades.

8.3. Event selection and resolutions

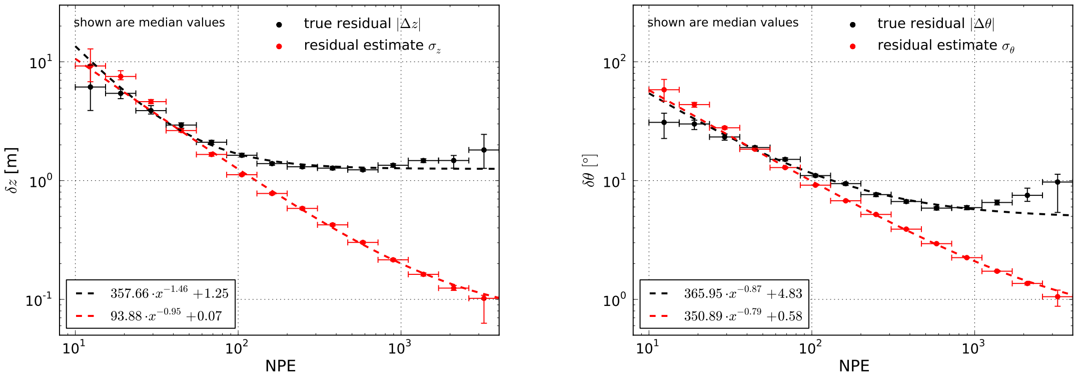

Figure 8.2.: Medians of true residuals (black points) and residual estimates (red points) for vertex depth z (left) and zenith angle θ (right) and fits as a function of observed event charge for ν_e CC events at V4 cut level in PINGU V36.

V4 cut level

Reconstructed events were required to satisfy the cuts “V4”. The first of these is based on a less computationally expensive reconstruction that is also performed and aims at reducing the rate of atmospheric background muons. The second, on the other hand, is based on Multinest itself. It requires the reconstructed vertex be within a right cylinder of 85 m radius from PINGU’s central axis, ranging from the the top of PINGU to the largest depth at which IceCube is instrumented below PINGU, and in addition that the reconstructed neutrino energy exceed 1 GeV. From this sample, only events reconstructed as upgoing, $-1 \leq \cos\theta < 0$, were then selected.

The covariance matrix of the ML estimators was calculated for each event assuming that a single cascade—with visible energy corresponding to the sum of that of the reconstructed cascade and the track energy—is a good solution to the reconstruction problem.

Similarly to the approach taken in chapter 7, Fig. 8.2 shows the median residuals in vertex depth and neutrino zenith angle dependent on the total event charge, together with the corresponding residual estimates. One can see that depth reconstruction is again dominated by a systematic component at charges larger than 200 PE, which is on the order of 1.25 m. Contrary to what is found in Fig. 7.2, which applied to events simulated using a different ice model and PINGU geometry V15, reconstructed with different spline tables, zenith angle resolution does not seem to improve anymore at charges above approximately 500 PE, which leads to a significant discrepancy to the predictions of the residual estimates in this region. Moreover, events were not discarded just because they had reached the upper time limit of the reconstruction, resulting in an increase of the residuals at the right end of the range, where most spline table lookups were necessary.

Quality cuts

When looking into the different residual estimates, it was found that the vertex position residual estimator, defined in the following as $\sigma_{\vec{x}} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$, showed correlation with the detector performance parameters of interest to the NMH measurement—zenith angle and energy resolutions—while for example the energy residual estimator did not. Also, $\sigma_{\vec{x}}$ and σ_θ were found to be almost completely positively correlated, with a correlation coefficient $\rho = 0.93$, as is shown in Fig. A.4 in the appendix.

Before $\sigma_{\vec{x}}$ was used to define cuts, the individual residual estimates for the vertex coordinates were corrected for the widths of their pull distributions (7.1). This is illustrated in Fig. 8.3 for the pull in x, P_x. In the left panel, one can see that the widths σ of the distributions increase rapidly with charge, reaching $\sigma \simeq 10$ in the highest-charge bin. In order to compensate for the deviation from $\sigma = 1$, each residual estimate was then scaled according to the actual width of the pull distribution in the corresponding charge range. The right panel of Fig. 8.3 shows the parameters of the pull distributions resulting from the correction of the individual estimates. The latter now describe the true residuals much better, leading to pull widths which are close to one over the whole range of charges.

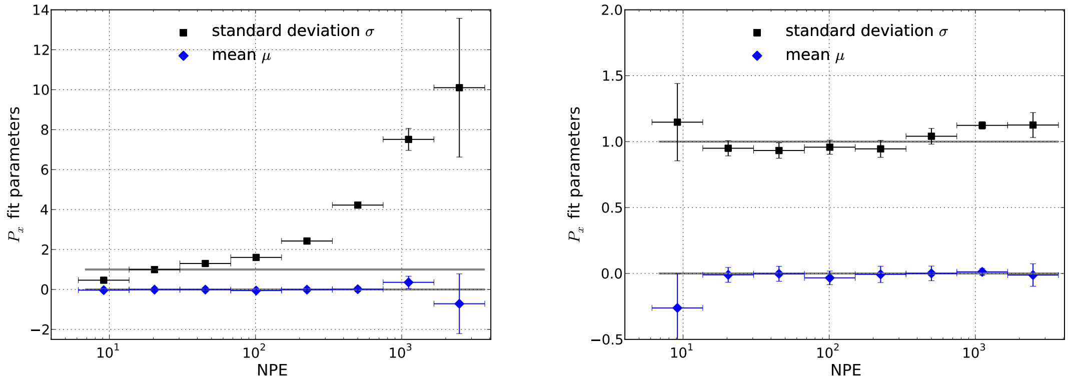

Figure 8.3.: Charge dependence of the pull distribution parameters for vertex depth z (left) and zenith angle θ (right). Shown are the means μ (in blue) and widths σ (in black) obtained from Gaussian fits. The gray horizontal lines correspond to the ideal values for an unbiased reconstruction whose residuals are on average perfectly predicted.

After the same procedure had been performed for the other two vertex coordinates, the "pull-corrected" position residual estimate accurately predicted the distance between the true and reconstructed vertex position. This can be seen in Fig. 8.4, where the medians are shown in dependence of the charge. Not only do they now scale very similarly, but they also agree well for events with the largest number of photoelectrons, where the vertex position resolution is on the order of 2.4 m.

This quantity was then used to define two straight cuts which split the original ν_e CC sample at V4 cut level into two subsamples each. As illustrated in Fig. 8.5, the first, "cut 1", was made at $\sigma_{\vec{x}} = 2.5$ m, while "cut 2" split the sample at $\sigma_{\vec{x}} = 4$ m. Approximately 30 %

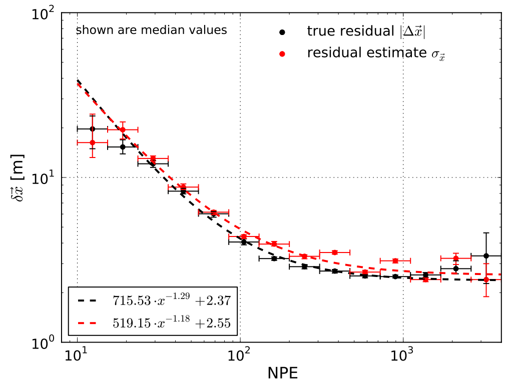

Figure 8.4.: Vertex position residuals and estimates after the latter had been corrected for the widths of the pull distributions in the three vertex coordinates.

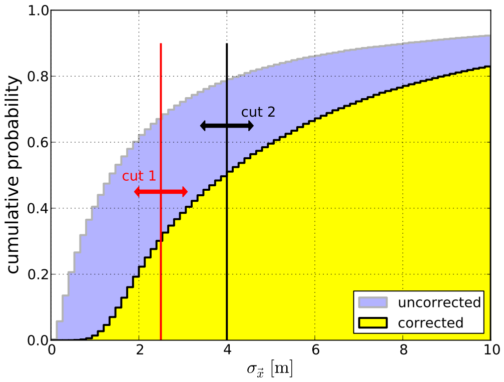

Figure 8.5.: Cumulative distributions of the corrected and uncorrected vertex position residual estimates. In addition, the definitions of the two cuts applied to the corrected estimates are shown. See text for details.

of all events satisfied cut 1, and just below 50 % of the events passed cut 2. Here, one can also identify the effect the pull width corrections had, which effectively scaled up the smallest estimates to account for the levelling-off of the resolution at high charges.

Fig. 8.6 shows the efficiencies of the two cuts on the corrected estimator $\sigma_{\vec{x}}$ as a function of true neutrino energy. At energies below 10 GeV, the first cut keeps between 5–30 % of the events in each bin, while the second cut at $\sigma_{\vec{x}} = 4\,\mathrm{m}$ retains between 20–50 %. For energies above 10 GeV, efficiencies are constant at around 35 % and 55 %, respectively.

Median vertex position, neutrino zenith angle and fractional neutrino energy resolutions are depicted in Fig. 8.7, as a function of true neutrino energy in the range 0–30 GeV, for events passing the quality cuts as well as for the original sample at V4 cut level.

The top plot shows that both cuts are very effective at selecting events whose vertex is reconstructed close to the true vertex. While the vertex resolution for the original sample is between 5 m and 6 m at $E_\nu = 5\,\mathrm{GeV}$, cut 2 reduces it to approximately 2.5 m and cut 1 somewhat further to 2 m.

As can be inferred from the central panel, in the original sample without further quality cuts zenith angle resolution improves from 20° at the lowest energies to approximately 8° at 20 GeV. The events passing the more relaxed quality cut have a zenith angle resolution of 15° at the low-energy end of the range shown, and are already constrained to within 6° of the original neutrino inclination for neutrino energies above 20 GeV. A further improvement of on the order 1° can be observed for cut 1.

Finally, the lowermost panel of Fig. 8.7 indicates that the cuts also result in samples with neutrino energy resolutions that are lowered by several percent, e.g. the median energy resolution of events passing cut 1 falls below 15 % at 15 GeV, compared to somewhat below 20 % for the V4 cut level sample.

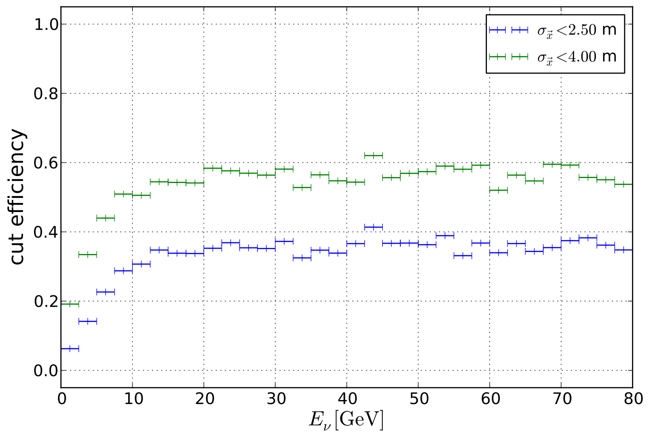

Figure 8.6.: Efficiencies for two straight cuts at 2.5 m and 4 m on the pull-corrected vertex resolution estimator $\sigma_{\vec{x}} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$ depending on true neutrino energy.

Before the, in total, five sets of events could be processed with the PaPA framework, $\cos\theta$ as well as energy resolutions were fit with double Gaussians in energy bins of 2 GeV width each in the range 0–20 GeV. Fig. A.5 in the appendix shows two fits to the distribution of the $\cos\theta$ residuals for true neutrino energies in the range 6–8 GeV, once for all events at V4 cut level, as well as for those that passed cut 2. There one can see how the width decreases by 20–30 % and that a bias toward small $\cos\theta$ is removed.

In addition, since the cuts that were applied exhibited energy-dependent efficiencies (Fig. 8.6), the effective area of the detector was redetermined anew for each sample. Since the event counts in each bin i in the $(E_\nu, \cos\theta)$ plane will then have the correct overall normalisation, this will permit adding the Fisher matrices of two independent samples, separated e.g. by the straight cuts that are described above.

In conclusion, five PINGU ν_e CC data sets with different properties were created. The first corresponds to all events that passed the basic event selection cuts V4, and is the one with the highest overall statistics. The next two were separated by cut 1 at $\sigma_{\vec{x}} = 2.5$ m, and contain approximately 30 % and 70 % of the number of events in the first set, respectively. Events that passed cut 1 are those with the best overall resolutions in vertex position, neutrino zenith angle and energy. Cut 2 at $\sigma_{\vec{x}} = 4$ m split the data set at V4 cut level in half, and events passing this cut also showed significant improvement in resolutions.

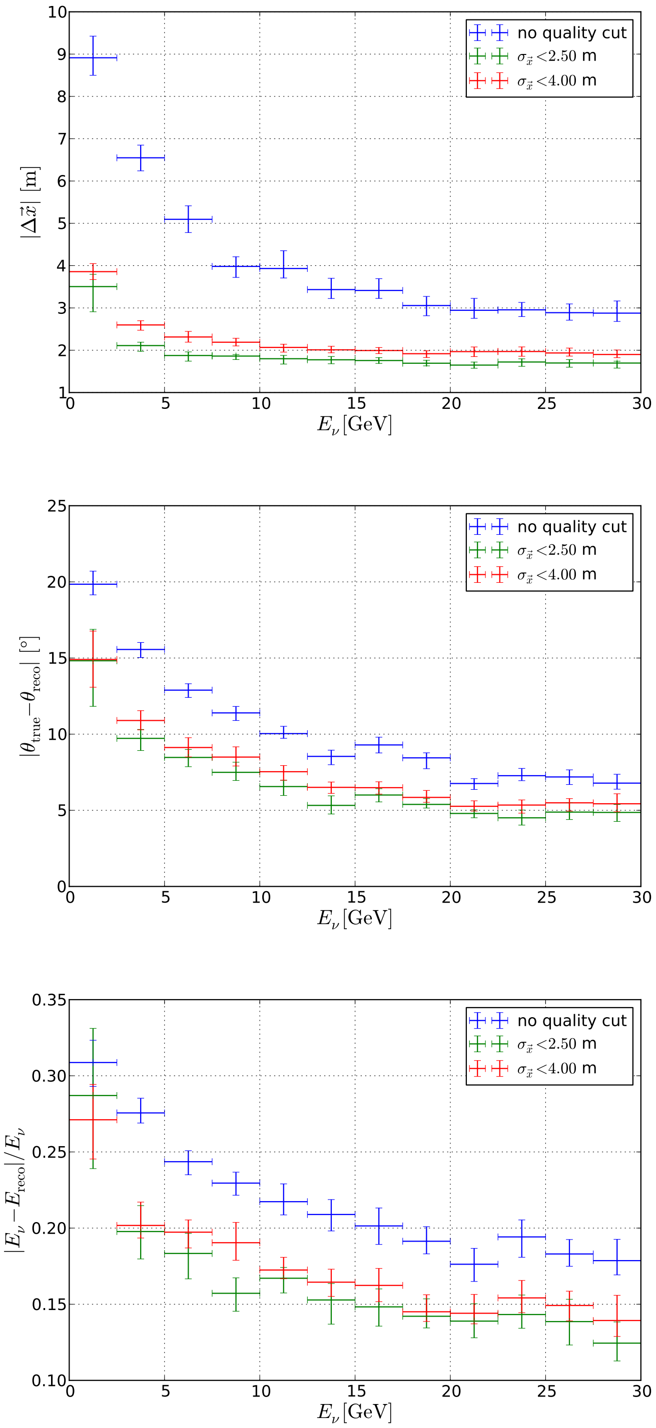

Figure 8.7.: Median resolutions as a function of true neutrino energy for the ν_e CC sample at V4 cut level and after the indicated cuts on the vertex position residual estimator have been performed.
Top: vertex position resolution. *Centre:* neutrino zenith angle resolution. *Bottom:* fractional neutrino energy resolution.

8.4. Systematics and significances

Apart from the mixing parameters already described in Sec. 8.1 when the PaPA framework was introduced, a set of additional parameters related to the detector response are included in the analysis. They allow for potential scaling errors on the reconstructed neutrino energy (s_E) and the energy dependence of the detector's effective area (s_{A_eff}), as well as for relative mismatches between neutrino- and anti-neutrino-nucleon cross sections ($n_{A_\text{eff},\nu/\bar{\nu}}$). Also, a scale factor related to the absolute normalisation of the atmospheric neutrino flux is taken into account, which, however, is only expected to affect the total number of events and could thus be replaced by a corresponding scaling of the detector livetime.

In general, the parameters that appear as systematics in the determination of the neutrino mass hierarchy are varied in finite steps around their fiducial values. The individual steps and fiducial values are detailed in Ref. [23]. It should be mentioned, however, that the fiducial value of the mass-squared difference Δm^2_{31} is taken as $2.46 \times 10^{-3}\,\text{eV}^2$, which corresponds to the normal hierarchy assumption.

The Fisher matrix (8.1) was evaluated for different detector livetimes. After its inversion one can extract the uncertainties of all systematic parameters as well as their correlations with the mass hierarchy parameter h.

Tab. 8.1 shows the results after 3 years of exposure for the ν_e CC sample at V4 cut level, i.e. with no further quality cut applied. In the left column the different systematic parameters are given, arranged according to the value of their squared correlation coefficient

$$c_{kl} = \frac{(F^{-1})_{kl}}{\sigma_k^\text{full}\sigma_l^\text{full}} \tag{8.2}$$

with the mass hierarchy parameter h. This is what is denoted as “impact” in the second column.

Priors are external constraints and correspond to the 1σ ranges for the mixing parameters. The *full* error on each parameter is given in the fourth column; it is the root of the sum of the squared *statistical* (fifth column) and *systematic* (sixth column) errors.

One can see that here the mixing angle θ_{23} has the largest impact on the mass hierarchy determination, with further major contributions stemming from a possible misscaling of the effective area with energy (or the tilt of the atmospheric neutrino spectrum) and the overall flux normalisation. The impacts of the latter are mainly due to an effective reduction in livetime and statistics, and more precise measurements of the flux will mitigate them [22]. Also, $\sin^2 2\theta_{23}$ is expected to have been determined more accurately, up to its octant, by the time PINGU starts taking data [88].

Tab. 8.2 shows the same quantities for the lowest-statistics sample which passed cut 1, also after 3 years of livetime. Here, impacts on the mass hierarchy parameter are in general significantly smaller, whereas statistical uncertainties are increased as expected from the lower event rates, especially at energies below 10 GeV, see Fig. 8.6. Since bin counts are smeared less strongly due to higher resolutions in E_ν and $\cos\theta$, degeneracies between actual mass hierarchy effects and uncertainties of the different systematic parameters can be resolved.

Parameter	Impact	Best Fit	Full	Stat.	Syst.	Prior
h	100.0	1.00	3.34×10^{-1}	2.38×10^{-1}	2.35×10^{-1}	free
θ_{23} [°]	31.1	38.6	1.14	6.65×10^{-1}	2.11	1.32
s_{A_eff} $[\frac{\mathrm{m}^2}{\mathrm{GeV}}]$	12.6	0.00	6.41×10^{-4}	4.29×10^{-4}	4.76×10^{-4}	free
flux norm.	11.7	0.00	2.24×10^{-2}	4.71×10^{-3}	2.20×10^{-2}	2.00×10^{-1}
θ_{13} [°]	5.0	8.93	4.64×10^{-1}	1.26	3.48	4.68×10^{-1}
Δm^2_{31} [eV^2]	4.3	0.00246	7.81×10^{-5}	2.01×10^{-4}	2.96×10^{-4}	8.00×10^{-5}
$n_{A_\mathrm{eff},\nu/\bar{\nu}}$	2.3	0.00	4.90×10^{-2}	1.10×10^{-2}	2.50×10^{-1}	5.00×10^{-2}
s_E	1.8	1.00	1.98×10^{-2}	8.08×10^{-2}	1.18×10^{-1}	2.00×10^{-2}

Table 8.1.: Uncertainties of the systematic parameters given the best-estimate detector configuration after 3 years of detector livetime for the ν_e CC event sample at V4 cut level, arranged according to their squared correlation with the NMH parameter h, denoted as “impact” here.

The median significance with which the wrong mass ordering can be excluded follows as the inverse of the full error on the mass hierarchy parameter h. In the absence of systematics, it is expected to scale as $\sqrt{t}$, where t is the detector livetime.

Fig. 8.8 shows the median significance as a function of detector livetime for the different samples. Here, the blue line is the significance that is obtained from whole set of ν_e CC events at V4 cut level, and the black line shows how it scales without systematics, i.e. if one inverts the diagonal element of the Fisher matrix corresponding to the hierarchy parameter. The measurement is already dominated by systematic effects after a fraction of one year’s worth of exposure time, so that one obtains a significance of, for example, 3σ after 3 years, compared to approximately 4.25σ if there were no systematics.

Comparing the significances for the samples divided by the cut at 2.5 m (dotted lines), on can see that the significance for the events that did not pass the cut grows faster in the beginning, but is more strongly impacted by systematics, so that after 10 years both samples almost yield the same value.

In the case of the 4 m cut, which split the sample in half, the sensitivity is larger as expected right from the beginning for the events which have better resolution, resulting in a difference of 0.5σ after 10 years of exposure.

The significances of the combined samples, corresponding to the orange and purple lines, were determined from the sum of the Fisher matrices of the individual samples. While after two years no improvement is seen for the combined significance of the two samples divided by the 2.5 m cut, the combination of the other two subsamples with each approximately 50 % of the number of original events leads to an increased sensitivity almost from the start.

Parameter	Impact	Best Fit	Full	Stat.	Syst.	Prior
h	100.0	1.00	4.53×10^{-1}	4.03×10^{-1}	2.06×10^{-1}	free
θ_{23} [°]	10.0	38.6	1.21	1.46	2.63	1.32
flux norm.	3.4	0.00	2.60×10^{-2}	1.25×10^{-2}	2.31×10^{-2}	2.00×10^{-1}
s_{A_eff} $[\frac{\text{m}^2}{\text{GeV}}]$	2.4	0.00	1.20×10^{-3}	8.41×10^{-4}	8.49×10^{-4}	free
$n_{A_{\text{eff}},\nu/\bar{\nu}}$	2.2	0.00	4.94×10^{-2}	3.17×10^{-2}	3.09×10^{-1}	5.00×10^{-2}
θ_{13} [°]	1.9	8.93	4.65×10^{-1}	2.36	3.79	4.68×10^{-1}
Δm^2_{31} [eV^2]	1.9	0.00246	7.82×10^{-5}	2.89×10^{-4}	2.36×10^{-4}	8.00×10^{-5}
s_E	0.8	1.00	1.98×10^{-2}	1.18×10^{-1}	1.08×10^{-1}	2.00×10^{-2}

Table 8.2.: Same as Tab. 8.1, but for the events that passed cut 1 ($\sim 30\,\%$ altogether).

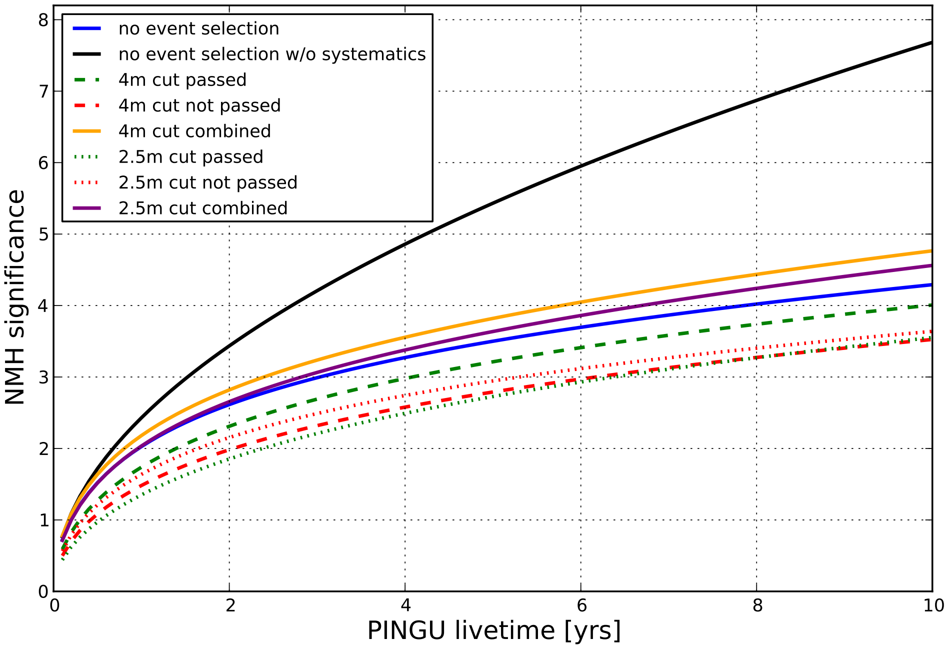

Figure 8.8.: Neutrino mass hierarchy significance as a function of PINGU livetime arising from the individual samples defined in Sec. 8.3 and the combination of those divided by cuts 1 & 2. In addition, the black line shows the $\sqrt{t}$ behaviour for the whole sample of ν_e CC events, assuming no impact from systematics.

One can also check how the uncertainties of the mixing parameters evolve with time. These are shown in Fig. 8.9 for Δm^2_{31} as well as θ_{23}. In the former case, better resolutions seem to have a rather strong impact on the uncertainties. The sample passing cut 2 (dashed green line), even though it only contains 50 % of the original events, yields reduced uncertainties after approximately 3 years of detector livetime with respect to the sample at V4 cut level. The overall uncertainties resulting from taking into account only the low-quality events (red dashed and dotted lines) are larger and decrease more slowly with time. Combining these with their higher-resolution counterparts, however, leads to a much faster reduction of the uncertainty on Δm^2_{31}.

For θ_{23}, the two lines corresponding to the samples divided by the cut at 2.5 m are switched. Also, the green dashed line now lies above the blue line representing the whole sample. It seems like here the loss in event rates at energies below 10 GeV, where the mass hierarchy induced matter effects of the oscillation probabilities are strongest (see Fig. 2.4), is detrimental to the determination of θ_{23}. The recombination of the two samples again, however, leads to stronger constraints on the value of the 2-3 mixing angle.

Finally, one should note that the overall scale of the uncertainties in this idealised analysis—which only considers the ν_e CC channel, assumes perfect flavour identification and has no contribution from NC events for which the energy resolution is worse—is large when compared to a more realistic scenario that considers both the cascade channel and the track channel, even with non-ideal flavour identification. Such an analysis performed for PINGU in its V15 configuration with parametrised resolutions yielded uncertainties of approximately $6.8 \times 10^{-5}\,\mathrm{eV}^2$ and 0.9° for Δm^2_{31} and θ_{23}, respectively, after one year of livetime [23].

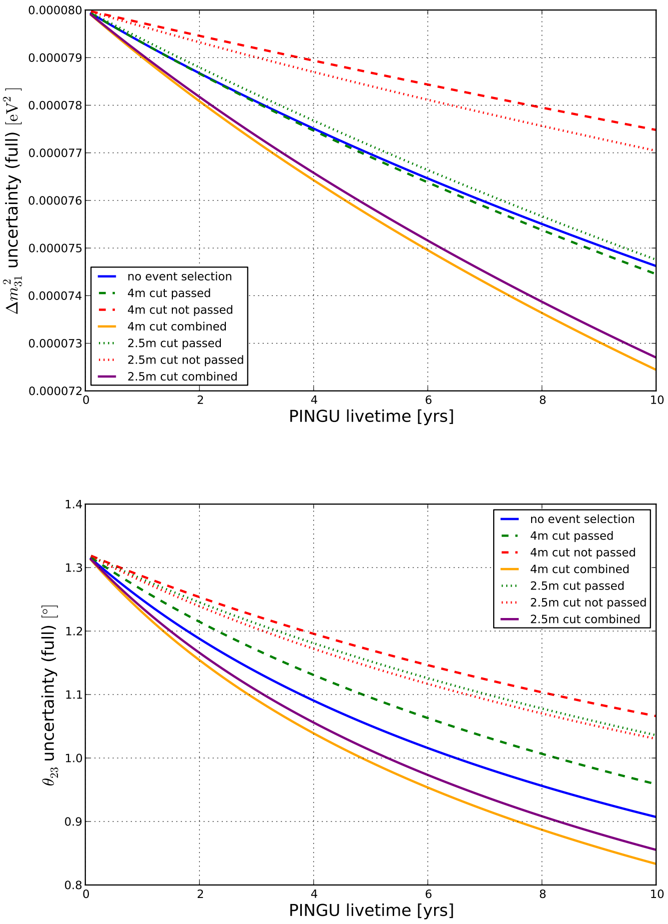

Figure 8.9.: Evolution of the full uncertainties of oscillation parameters with PINGU livetime for the different ν_e CC data sets. *Top:* Δm^2_{31}. *Bottom:* θ_{23}. In each case, the blue line shows the full sample that did not undergo further quality cuts.

9. Summary and Outlook

Event reconstruction in PINGU is of universal nature: the visible energy of a hadronic cascade arising from all kinds of charged-current as well as neutral-current neutrino interactions in the GeV regime can be significantly smaller than that of an electromagnetic cascade of the same energy. For this reason, the hybrid reconstruction approach Multinest is employed, which assumes that the underlying event is a charged-current muon neutrino interaction.

This work dealt with the reconstruction of cascade-like events in PINGU. These have been shown to be important in that they can contribute greatly to the correct determination of the neutrino mass hierarchy, which is one of the main physics goals of PINGU.

First, it was shown that, in the vicinity of its maximum, a likelihood function can provide information about the intrinsic uncertainty of estimators. An existing approach aiming at predicting the latter in the case of cascades, which determines the covariance matrix of event parameters from the expectation value of the second derivatives of the log-likelihood function, was presented. Then, a possible adaptation of this approach to the complex event hypothesis employed by a hybrid reconstruction such as Multinest was given.

In the main part of this work, criteria which the likelihood is expected to satisfy when it represents a suitable description of the actual detector response were introduced. The hybrid residual estimator was applied a-posteriori to simulated ν_e CC events reconstructed with Multinest, and the strongest deviations were found at highest charges, where consistency between simulation and assumptions in the likelihood function is most important.

Two possible sources of error, wrong assumptions about the Cherenkov amplitudes or about photon delay times, were disentangled by not taking into account photon arrival times in reconstruction. The findings suggested discrepancies between tabulated light yields and light yields resulting from a full particle propagation, which could in turn lead to systematic disagreement between the maximum likelihood estimates and the values of the true event parameters.

For a simplified simulation where photons were generated using spline tables, systematic influences especially on vertex reconstruction at highest charges were much reduced. As a result, the predictions of the likelihood-based residual estimators were relatively accurate over the whole charge range. Also here, however, the likelihood formulation favoured event parameters that were further away from the true values than expected, especially when photon timing was employed.

After simulation had been further deconstructed by removing all non-Poissonian elements such as uncertainties in hadronic light output, irregular PMT pulses or electronics effects, a reconstruction that only considered detected photon numbers lead to statistical agreement between the true and reconstructed parameter values over the whole range of charges.

In the end, a different photon timing scheme, making use of equally spaced bins, was devised. The smallest bin width of 5 ns seemed most promising, but was also most computationally expensive.

In the second part, the likelihood-based residual estimator was used to perform an event selection of ν_e CC events simulated with PINGU geometry V36, in order to estimate the implications of different resolutions for the neutrino mass hierarchy sensitivity. The vertex position residual estimator was found to have discriminating power in neutrino zenith angle and energy resolutions. After pull distribution widths deviating from one had been corrected for, two cuts were defined and the resulting resolutions were parametrised for use with the *Parametric PINGU Analysis*. Results implied that better resolutions resolved degeneracies between systematic parameters such as the mixing angle θ_{23} and the mass hierarchy. Furthermore, it could be shown that both the mass hierarchy significance as well as uncertainties of the mixing parameters can profit from a selection of well-reconstructed events which are recombined with the rest of the sample.

There are several aspects which can be taken up in the future. For one, it would be interesting to find out whether improved spline tables or different photon timing schemes can mitigate some of the problems encountered. The extension of the cascade residual estimator presented in this work could be applied to tau or muon induced events, and an extended mass hierarchy sensitivity analysis could be performed by taking into account all channels. In the case of muon neutrino charged-current scattering, its hypothesis would actually agree with the underlying event signature. It should also allow for examining correlations between different estimators; in principle, Multinest could also be able to provide flavour identification possibilities if degeneracies between the length of the track and the energy of the cascade can be resolved.

A. Appendix

PINGU V15 total CC neutrino energy and event charge

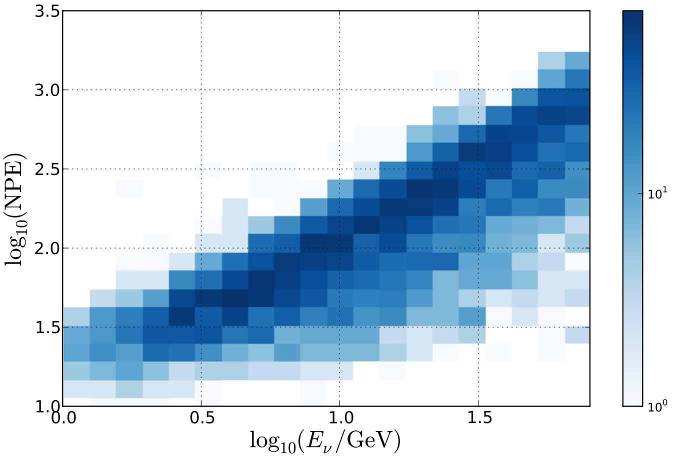

Figure A.1.: Total event charge as a function of neutrino energy for ν_e CC MC events that reconstructed within the fiducial volume of PINGU V15.

Pull distributions for different charges

Default reconstruction scenario - vertex depth

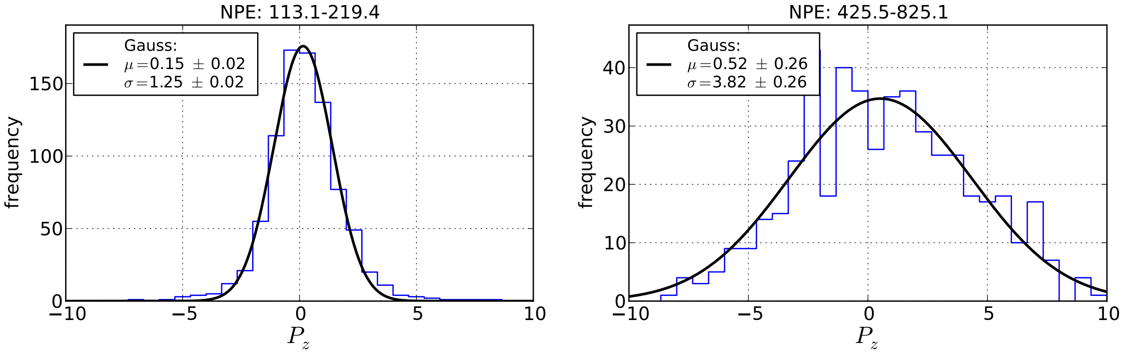

Figure A.2.: Gaussian fits to the pull distributions in vertex depth z for two different charge ranges for the default Multinest reconstruction scenario.

Log-likelihood ratio for idealised simulation-reconstruction loop with default timing

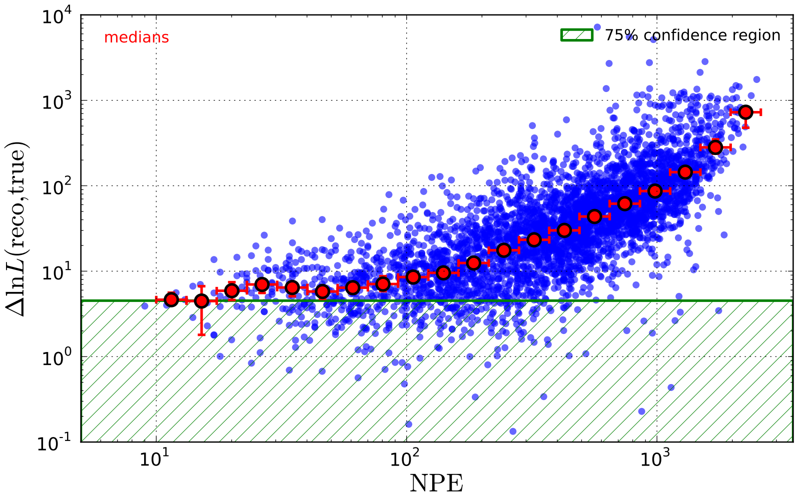

Figure A.3.: Scaling of the log-likelihood ratio of the idealised simulation-reconstruction loop with default photon timing of 1 photon per bin.

Neutrino mass hierarchy analysis

Estimator correlation

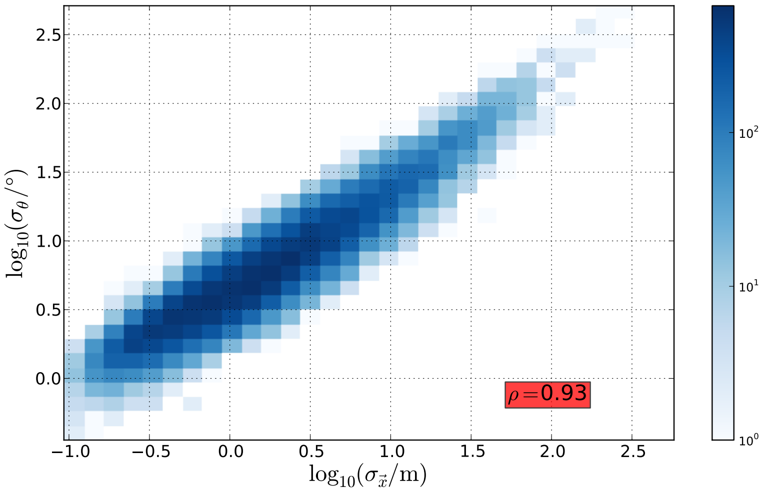

Figure A.4.: Correlation between the vertex position residual estimator $\sigma_{\vec{x}}$ and the zenith angle residual estimator σ_θ for ν_e CC events in PINGU V36

Parametrised resolutions for use with PaPA

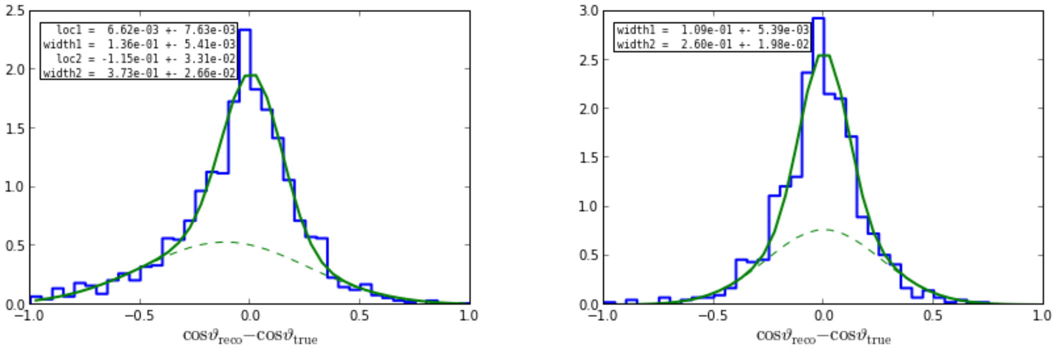

Figure A.5.: Double-Gaussian fits of cosine zenith resolutions for true neutrino energies in the range 6–8 GeV. The left panel shows the distribution of the residuals for the whole ν_e CC sample at V4 cut level, whereas the sample shown in the right only consists of events that in addition passed the 4 m cut on the corrected vertex position resolution estimator. The relative strength of the second Gaussian amounts to 0.5. Furthermore, in the right panel, the means of both Gaussians are fixed to zero.

Bibliography

[1] M. G. Aartsen et al. "Letter of Intent: The Precision IceCube Next Generation Upgrade (PINGU)". In: *ArXiv e-prints* (Jan. 2014). arXiv: `1401.2046v1 [physics.ins-det]`.

[2] M. G. Aartsen et al. "Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data". In: *ArXiv e-prints* (2014). arXiv: `1410.7227 [hep-ex]`.

[3] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: `10.1103/PhysRev.105.1413`. URL: `http://link.aps.org/doi/10.1103/PhysRev.105.1413`.

[4] T. D. Lee and C. N. Yang. "Parity Nonconservation and a Two-Component Theory of the Neutrino". In: *Phys. Rev.* 105 (5 Mar. 1957), pp. 1671–1675. DOI: `10.1103/PhysRev.105.1671`. URL: `http://link.aps.org/doi/10.1103/PhysRev.105.1671`.

[5] R. P. Feynman and M. Gell-Mann. "Theory of the Fermi Interaction". In: *Phys. Rev.* 109 (1 Jan. 1958), pp. 193–198. DOI: `10.1103/PhysRev.109.193`. URL: `http://link.aps.org/doi/10.1103/PhysRev.109.193`.

[6] N. G. Cooper [ed.] *Los Alamos Science, Number 25: Celebrating the Neutrino*. 1997.

[7] R. Davis, D. S. Harmer, and K. C. Hoffman. "Search for Neutrinos from the Sun". In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: `10.1103/PhysRevLett.20.1205`. URL: `http://link.aps.org/doi/10.1103/PhysRevLett.20.1205`.

[8] K. A. Olive. "Review of Particle Physics". In: *Chin. Phys. C* 38 (2014), p. 090001.

[9] V.N. Aseev et al. "An upper limit on electron antineutrino mass from Troitsk experiment". In: *Phys. Rev.* D84 (2011), p. 112003. DOI: `10.1103/PhysRevD.84.112003`. arXiv: `1108.5034 [hep-ex]`.

[10] Ch. Kraus et al. "Final results from phase II of the Mainz neutrino mass search in tritium beta decay". In: *Eur. Phys. J.* C40 (2005), pp. 447–468. DOI: `10.1140/epjc/s2005-02139-7`. arXiv: `hep-ex/0412056 [hep-ex]`.

[11] K.N. Abazajian et al. "Cosmological and Astrophysical Neutrino Mass Measurements". In: *Astropart. Phys.* 35 (2011), pp. 177–184. DOI: `10.1016/j.astropartphys.2011.07.002`. arXiv: `1103.5083 [astro-ph.CO]`.

[12] P.A.R. Ade et al. "Planck 2013 results. XVI. Cosmological parameters". In: *Astron. Astrophys.* (2014). DOI: `10.1051/0004-6361/201321591`. arXiv: `1303.5076 [astro-ph.CO]`.

[13] K. Zuber. *Neutrino Physics*. Bristol: Institute of Physics Publishing, 2004.

[14] A. Schukraft. “A view of prompt atmospheric neutrinos with IceCube”. In: *Nucl. Phys. Proc. Suppl.* 237-238 (2013), pp. 266–268. DOI: `10.1016/j.nuclphysbps.2013.04.105`. arXiv: `1302.0127 [astro-ph.HE]`.

[15] M. G. Aartsen et al. “Development of a General Analysis and Unfolding Scheme and its Application to Measure the Energy Spectrum of Atmospheric Neutrinos with IceCube”. In: *ArXiv e-prints* (Sept. 2014). arXiv: `1409.4535v1 [astro-ph.HE]`.

[16] Z. Maki, M. Nakagawa, and S. Sakata. “Remarks on the Unified Model of Elementary Particles”. In: *Progress of Theoretical Physics* 28.5 (1962), pp. 870–880. DOI: `10.1143/PTP.28.870`. eprint: `http://ptp.oxfordjournals.org/content/28/5/870.full.pdf+html`. URL: `http://ptp.oxfordjournals.org/content/28/5/870.abstract`.

[17] C. Berger. *Elementarteilchenphysik: Von den Grundlagen zu den modernen Experimenten*. 2nd edition. Berlin Heidelberg New York: Springer Verlag, 2006.

[18] E. K. Akhmedov. “Neutrino oscillations beyond two flavors”. In: *Nucl. Phys. Proc. Suppl.* 118 (2003), pp. 245–254. DOI: `10.1016/S0920-5632(03)01322-7`. arXiv: `hep-ph/0207342 [hep-ph]`.

[19] A. M. Dziewonski and D. L. Anderson. “Preliminary reference Earth model”. In: *Physics of the Earth and Planetary Interiors* 25.4 (1981), pp. 297–356.

[20] E. K. Akhmedov et al. “Atmospheric neutrinos at Super-Kamiokande and parametric resonance in neutrino oscillations”. In: *Nucl. Phys.* B542 (1999), pp. 3–30. DOI: `10.1016/S0550-3213(98)00825-6`. arXiv: `hep-ph/9808270 [hep-ph]`.

[21] A. Kouchner. “Next-generation atmospheric neutrino experiments”. In: *Physics of the Dark Universe* 4 (2014), pp. 60–74. ISSN: 2212-6864. DOI: `http://dx.doi.org/10.1016/j.dark.2014.09.001`. URL: `http://www.sciencedirect.com/science/article/pii/S2212686414000259`.

[22] E. K. Akhmedov, S. Razzaque, and A. Y. Smirnov. “Mass hierarchy, 2-3 mixing and CP-phase with Huge Atmospheric Neutrino Detectors”. In: *JHEP* 1302 (2013), p. 082. DOI: `10.1007/JHEP02(2013)082, 10.1007/JHEP07(2013)026`. arXiv: `1205.7071 [hep-ph]`.

[23] L. Schulte, M. Kowalski, and S. Böser. *Evaluating PINGU Sensitivity to the Neutrino Mass Hierarchy using the Fisher Information Matrix*. May 2014. URL: `http://internal.icecube.wisc.edu/reports/data/icecube/2014/05/001/icecube_201405001_v2.pdf`.

[24] J. A. Formaggio and G. P. Zeller. “From eV to EeV: Neutrino cross sections across energy scales”. In: *Rev. Mod. Phys.* 84 (3 Sept. 2012), pp. 1307–1341. DOI: `10.1103/RevModPhys.84.1307`. URL: `http://link.aps.org/doi/10.1103/RevModPhys.84.1307`.

[25] J. Beringer et al. “Review of Particle Physics”. In: *Phys. Rev. D* 86 (1 July 2012), p. 010001. DOI: `10.1103/PhysRevD.86.010001`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.86.010001`.

[26] D. Groom. *Particle Data Group Website. Atomic and Nuclear Properties of Materials, for more than 300 materials.* 2014. URL: `http://pdg.lbl.gov/2014/AtomicNuclearProperties` (visited on 11/27/2014).

[27] L. Rädel and C. Wiebusch. “Calculation of the Cherenkov light yield from electromagnetic cascades in ice with Geant4”. In: *Astroparticle Physics* 44 (Apr. 2013), pp. 102–113. DOI: `10.1016/j.astropartphys.2013.01.015`. arXiv: `1210.5140 [astro-ph.IM]`.

[28] M. Ackermann et al. “Optical properties of deep glacial ice at the South Pole”. In: *Journal of Geophysical Research (Atmospheres)* 111, D13203 (July 2006). DOI: `10.1029/2005JD006687`.

[29] J. V. Jelley. “Cerenkov radiation and its applications”. In: *British Journal of Applied Physics* 6.7 (1955), pp. 227–232. URL: `http://stacks.iop.org/0508-3443/6/i=7/a=301`.

[30] W. Heitler. *Quantum Theory of Radiation.* 2nd edition. Oxford: Oxford University Press, 1944.

[31] J. F. Carlson and J. R. Oppenheimer. In: *Phys. Rev.* 51 (1937).

[32] C. Wiebusch. “The Detection of Faint Light in Deep Underwater Neutrino Telescopes”. PhD thesis. RWTH Aachen, 1995.

[33] M. Kowalski. *On the Cherenkov light emission of hadronic and electro-magnetic cascades.* Aug. 12, 2002. URL: `http://internal.icecube.wisc.edu/reports/amanda/data/20020803-track.pdf`.

[34] T.A. Gabriel et al. “Energy dependence of hadronic activity”. In: *Nuclear Instruments and Methods in Physics Research A* 338 (1994), pp. 336–347. ISSN: 0168-9002. DOI: `http://dx.doi.org/10.1016/0168-9002(94)91317-X`. URL: `http://www.sciencedirect.com/science/article/pii/016890029491317X`.

[35] D.E. Groom. “Energy flow in a hadronic cascade: Application to hadron calorimetry”. In: *Nuclear Instruments and Methods in Physics Research A* 572.2 (2007), pp. 633–653. ISSN: 0168-9002. DOI: `http://dx.doi.org/10.1016/j.nima.2006.11.070`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900206023965`.

[36] M. Kowalski. “Search for Neutrino-Induced Cascades with the AMANDA-II Detector”. PhD thesis. Humboldt-Universität zu Berlin, 2003.

[37] IceCube Collaboration. *IceCube Preliminary Design Document.* 2001. URL: `http://icecube.wisc.edu/icecube/static/reports/IceCubeDesignDoc.pdf`.

[38] R. Abbasi et al. “The design and performance of IceCube DeepCore”. In: *Astropart. Phys.* 35.10 (2012), pp. 615–624. ISSN: 0927-6505. DOI: http://dx.doi.org/10.1016/j.astropartphys.2012.01.004. URL: http://www.sciencedirect.com/science/article/pii/S0927650512000254.

[39] Hamamatsu Photonics. *Photomultiplier Tubes. Basics and Applications*. Edition 3a. 2007.

[40] *The IceCube coordinate system*. URL: https://wiki.icecube.wisc.edu/index.php/Coordinate_system (visited on 11/26/2014).

[41] R. Abbasi et al. “The IceCube data acquisition system: Signal capture, digitization, and timestamping”. In: *Nuclear Instruments and Methods in Physics Research A* 601.3 (2009), pp. 294–316. ISSN: 0168-9002. DOI: http://dx.doi.org/10.1016/j.nima.2009.01.001. arXiv: 0810.4930 [physics.ins-det].

[42] G. Mie. “Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen”. In: *Annalen der Physik* 330.3 (1908), pp. 377–445. ISSN: 1521-3889. DOI: 10.1002/andp.19083300302. URL: http://dx.doi.org/10.1002/andp.19083300302.

[43] P. Askebjer et al. “Optical properties of the South Pole ice at depths between 0.8-km and 1-km”. In: *Science* 267 (1995), pp. 1147–1150. DOI: 10.1126/science.267.5201.1147. arXiv: astro-ph/9412028 [astro-ph].

[44] A. Ishimaru. *Wave Propagation and Scattering in Random Media: Single scattering and transport theory*. Wave Propagation and Scattering in Random Media. Academic Press, 1978. URL: http://books.google.de/books?id=tLsRAQAAIAAJ.

[45] M. G. Aartsen et al. “Measurement of South Pole ice transparency with the IceCube LED calibration system”. In: *Nuclear Instruments and Methods in Physics Research A* 711 (May 2013), pp. 73–89. arXiv: 1301.5361 [astro-ph.IM].

[46] P. Askebjer et al. “UV and optical light transmission properties in deep ice at the South Pole”. In: *Geophysical Research Letters* 24.11 (1997), pp. 1355–1358. ISSN: 1944-8007. DOI: 10.1029/97GL01246. URL: http://dx.doi.org/10.1029/97GL01246.

[47] S. L. Miller. “Clathrate Hydrates of Air in Antarctic Ice”. In: *Science* 165 (Aug. 1969), pp. 489–490.

[48] T. Ushida et al. “Refractive-index measurements of natural air hydrate crystals in an Antarctic ice sheet”. In: *Appl. Opt.* 34 (Sept. 1995), pp. 5746–5749.

[49] P. B. Price, K. Woschnagg, and D. Chirkin. “Age vs depth of glacial ice at South Pole”. In: *Geophysical Research Letters* 27.14 (2000), pp. 2129–2132. ISSN: 1944-8007. DOI: 10.1029/2000GL011351. URL: http://dx.doi.org/10.1029/2000GL011351.

[50] *AHA ice model*. URL: https://wiki.icecube.wisc.edu/index.php/AHA_ice_model (visited on 11/17/2014).

[51] C. Andreopoulos et al. "The GENIE neutrino Monte Carlo generator". In: *Nuclear Instruments and Methods in Physics Research A* 614 (Feb. 2010), pp. 87–104. DOI: 10.1016/j.nima.2009.12.009. arXiv: 0905.2517 [hep-ph].

[52] M.G. Aartsen et al. "Energy Reconstruction Methods in the IceCube Neutrino Telescope". In: *JINST* 9 (2014), P03009. DOI: 10.1088/1748-0221/9/03/P03009. arXiv: 1311.4767 [physics.ins-det].

[53] *clsim module documentation*. URL: http://software.icecube.wisc.edu/simulation_trunk/projects/clsim (visited on 08/26/2014).

[54] *Vuvuzela source code*. URL: http://code.icecube.wisc.edu/svn/projects/vuvuzela/ (visited on 08/26/2014).

[55] *PMTResponseSimulator module documentation*. URL: http://software.icecube.wisc.edu/simulation_trunk/projects/DOMLauncher/PMTRes.html (visited on 08/26/2014).

[56] R. Abbasi et al. "Calibration and characterization of the IceCube photomultiplier tube". In: *Nuclear Instruments and Methods in Physics Research A* 618 (June 2010), pp. 139–152. DOI: 10.1016/j.nima.2010.03.102. arXiv: 1002.2442 [astro-ph.IM].

[57] *SPE Time Distribution*. URL: https://wiki.icecube.wisc.edu/index.php/SPE_Time_Distribution (visited on 08/26/2014).

[58] *Prepulse Data*. URL: https://wiki.icecube.wisc.edu/index.php/Prepulse_Data (visited on 08/26/2014).

[59] *Late Pulse Data*. URL: https://wiki.icecube.wisc.edu/index.php/Late_Pulse_data (visited on 08/26/2014).

[60] *Afterpulse Data*. URL: https://wiki.icecube.wisc.edu/index.php/Afterpulse_Data (visited on 08/26/2014).

[61] *DOMLauncher module documentation*. URL: http://software.icecube.wisc.edu/simulation_trunk/projects/DOMLauncher/DOML.html (visited on 08/26/2014).

[62] *WaveCalibrator module documentation*. URL: http://software.icecube.wisc.edu/simulation_trunk/projects/WaveCalibrator/index.html (visited on 08/26/2014).

[63] *Digitizer baseline calibration*. URL: https://wiki.icecube.wisc.edu/index.php/Digitizer_baseline_calibration (visited on 08/27/2014).

[64] *Droop correction*. URL: https://wiki.icecube.wisc.edu/index.php/Droop_correction (visited on 08/27/2014).

[65] C. Wendt. *Droop Correction - Dual tau Model*. 2006. URL: http://docushare.icecube.wisc.edu/docushare/dsweb/Get/Document-30244/Droop-dual-tau-Zeuthen2006_wendt.pdf (visited on 08/27/2014).

[66] C. Roucelle. *Documentation for the DOMcalibrator module (offline-software release V01-11-02)*. Sept. 10, 2007.

[67] *Wavedeform module documentation*. URL: http://software.icecube.wisc.edu/simulation_trunk/projects/wavedeform (visited on 08/26/2014).

[68] N. Whitehorn. "A Search for High-Energy Neutrino Emission from Gamma-Ray Bursts". PhD thesis. University of Wisconsin-Madison, 2012.

[69] M. Usner. "Search for Neutrino-induced Cascades in IceCube". Master's thesis. Rheinische Friedrich-Wilhelms-Universität Bonn, 2012.

[70] G. Cowan. *Statistical Data Analysis*. 1st edition. Oxford New York: Oxford University Press, 1998.

[71] D. S. Sivia and J. Skilling. *Data Analysis: A Bayesian Tutorial*. 2nd edition. Oxford New York: Oxford University Press, 2006.

[72] S. Brandt. *Data Analysis*. 3rd edition. New York Berlin Heidelberg: Springer-Verlag, 1997.

[73] F. James. *Statistical Methods in Experimental Physics*. 2nd edition. World Scientific, 2006.

[74] J. v. Santen. "Markov-Chain Monte-Carlo Reconstruction for cascade-like events in IceCube". Diploma thesis. Humboldt-Universität zu Berlin, 2010.

[75] *Millipede source code*. URL: http://code.icecube.wisc.edu/svn/projects/millipede/ (visited on 08/26/2014).

[76] J. Lundberg et al. "Light tracking through ice and water - Scattering and absorption in heterogeneous media with Photonics". In: *Nuclear Instruments and Methods in Physics Research Section A* 581.3 (2007), pp. 619–631. DOI: 10.1016/j.nima.2007.07.143. arXiv: 0702108v2 [astro-ph].

[77] N. Whitehorn, J. van Santen, and S. Lafebe. "Penalized Splines for Smooth Representation of High-dimensional Monte Carlo Datasets". In: *Computer Physics Communications* 184 (2013), pp. 2214–2220. DOI: 10.1016/j.cpc.2014.04.008. arXiv: 1301.2184.

[78] F. Feroz, M.P. Hobson, and M. Bridges. "MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics". In: *Mon. Not. Roy. Astron. Soc.* 398 (2009), pp. 1601–1614. DOI: 10.1111/j.1365-2966.2009.14548.x. arXiv: 0809.3437 [astro-ph].

[79] F. Feroz et al. "Importance Nested Sampling and the MultiNest Algorithm". In: *ArXiv e-prints* (Jan. 2014). arXiv: 1306.2144v2 [astro-ph.IM].

[80] *Multinest source code*. URL: http://code.icecube.wisc.edu/projects/icecube/browser/IceCube/sandbox/multinest_icetray (visited on 08/26/2014).

[81] K. Wiebe. *Resca source code*. URL: http://code.icecube.wisc.edu/projects/icecube/browser/IceCube/sandbox/kwiebe/resca (visited on 08/26/2014).

[82] W. Winter. “Neutrino mass hierarchy determination with IceCube-PINGU”. In: *Phys. Rev.* D88.1 (2013), p. 013013. DOI: 10.1103/PhysRevD.88.013013. arXiv: 1305.5539 [hep-ph].

[83] T. Neunhöffer. “Die Entwicklung eines neuen Verfahrens zur Suche nach kosmischen Neutrino-Punktquellen mit dem AMANDA-Neutrinoteleskop”. PhD thesis. Johannes Gutenberg-Universität Mainz, 2003.

[84] L. Demortier and L. Lyons. *"Everything you always wanted to know about pulls"*. CDF note 5776. Aug. 2002. URL: http://physics.rockefeller.edu/luc/technical_reports/cdf5776_pulls.pdf.

[85] E. Middell. “Reconstruction of Cascade-Like Events in IceCube”. Diploma thesis. Humboldt-Universität zu Berlin, 2008.

[86] D. Góra et al. *Studies on Millipede*. Apr. 29, 2014. URL: http://internal.icecube.wisc.edu/reports/data/icecube/2014/04/001/icecube_201404001_v1.pdf.

[87] D. Góra. Personal communication.

[88] M. Blennow et al. “Quantifying the sensitivity of oscillation experiments to the neutrino mass ordering”. In: *ArXiv e-prints* (2013). arXiv: arXiv:1311.1822v1 [hep-ph].

List of Figures

List of Tables

Danksagung

Ohne die Unterstützung einer Vielzahl von Personen wäre diese Arbeit nicht möglich gewesen.

Zunächst möchte ich deshalb Herrn Prof. Dr. Marek Kowalski meine Dankbarkeit dafür aussprechen, dass er mir die Gelegenheit gab, in einer solch vielfältigen Gruppe und einem solch dynamischen Umfeld zu forschen. Ich bin sehr froh über die Erfahrungen, die ich etwa auf dem IceCube-Bootcamp oder dem Kollaborationsmeeting in Genf machen durfte.

Allen ehemaligen Mitgliedern sowie "Noch-Bonnern" unserer Arbeitsgruppe gilt mein Dank für die schöne Zusammenarbeit. Ein herzliches Dankeschön richten möchte ich im Speziellen an Lukas Schulte für all seine Hilfe bei Fragen zu PINGU und PaPA, an Prof. Dr. Sebastian Böser und Marcel Usner für ihre alles andere als selbstverständliche Unterstützung in vielerlei Hinsicht, zuletzt sogar aus der Distanz, sowie an Markus Voge, ohne dessen Einsatz ich bei einigen Softwarefragen aufgeschmissen gewesen wäre.

Nicht vergessen möchte ich an dieser Stelle auch Herrn Prof. Dr. Jochen Dingfelder für seine Bereitschaft, das Zweitgutachten zu übernehmen, herzlich zu danken.

Zu guter Letzt geht noch ein riesengroßes Dankeschön an meine Familie, ohne deren Verständnis und kontinuierliche Unterstützung mein Studium gar nicht erst möglich gewesen wäre.

Erklärung

Ich versichere, dass ich die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Bonn, den 8. Dezember 2014 Thomas Ehrhardt

Declaration

I hereby declare that this thesis was formulated by myself and that no sources or tools other than those cited were used.

Bonn, December the 8th 2014 Thomas Ehrhardt