Cascade type identification in IceCube and an application in a search for new physics
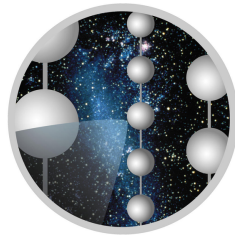
ICECUBE

Anna Katharina Steuer
ETAP
Institut für Physik
Staudingerweg 7
Johannes Gutenberg-Universität Mainz
D-55128 Mainz, Germany
ansteuer@uni-mainz.de

Abstract

In 2013, the first detection of an astrophysical neutrino flux by the IceCube neutrino telescope marked a tremendous breakthrough in the field of neutrino astronomy. Since then, dedicated analyses try to extract information on the origin of the cosmic messengers from the accumulating data. Efforts to trace the neutrino flux back to a dominant source class have not yet succeeded, leaving many questions on the nature of astrophysical particle accelerators unanswered and providing room for alternative explanations for the origin of IceCube's high energy events.

Besides the spatial distribution, the flavor composition of the astrophysical neutrino flux plays a key role in unraveling its production mechanisms. While the long *tracks* of Cherenkov light generated in muon neutrino charged current interactions can be well discriminated from the spherical shower - or *cascade* - pattern produced in all other neutrino interactions, an identification of the cascade-inducing neutrino flavor and interaction channel has yet to be accomplished. In this thesis, a method utilizing the delayed light from shower-internal neutron capture events - the so-called *neutron echo* - has been developed, experimentally implemented and applied to determine the shower type composition of a high energy cascade sample. This is the first time it has ever been attempted to classify the shower type of high energy cascades in IceCube or any other neutrino telescope and the results of this work show great promise.

Since the nature of the showers contained in IceCube's well-known *High Energy Starting Event* (HESE) sample has not been assessed before, an exotic origin could not be excluded. In this work, the newly developed method is used to probe two *boosted dark matter* scenarios, which predict an either solely neutral current- or electron neutrino charged current-like cascade sample. Although the so far small sample size limits the power of the analysis, a Bayesian evaluation of the results already disfavors the probed models at more than $90\,\%$ credible level.

Remember to look up at the stars and not down at your feet. ... It matters that you don't just give up.

- Stephen Hawking, ✝ March 14, 2018.

Contents

List of Figures

List of Tables

0.1 Introduction

Modern astronomy faces a wealth of puzzles yet to be solved. Although violent astrophysical events like supernovae, gamma-ray bursts and blazars have been detected, very little is known of the underlying physics processes that let these objects accelerate particles up to the highest energies. An even bigger mystery is the universe's dark matter content, from which we know that it exists, but whose nature has not yet been decrypted. Theory provides a variety of models to describe these phenomena, but experimental data are needed to either verify or reject them.

In exploring the small-scale physics of large-scale astrophysical objects and phenomena, neutrino astronomy provides formidable input. Neutrinos - the *cosmic messengers* - carry information on the physics processes at their production site through the universe without being deflected or absorbed. The IceCube neutrino telescope, situated at the geographic South Pole, monitors a cubic kilometer of Antarctic ice with 5160 optical modules to catch the light signature of neutrino interactions. The experiment facilitated the first discovery of an astrophysical neutrino flux in 2013. The energy spectrum, spatial distribution and flavor composition of astrophysical neutrinos entails valuable information on the conditions at their production sites.

So far, the main restriction in the determination of the astrophysical flavor composition is the indistinguishable event topology of electron and tau neutrinos. While the muon produced in a muon neutrino charged current interaction induces a well identifiable *track* of Cherenkov light, the intrinsic differences of the particle showers induced in all other neutrino interactions are extremely difficult to resolve in a coarse detector like IceCube. They appear as an almost spherical Cherenkov light pattern called *cascade*. Only above a few hundred TeV deposited energy, the tau neutrino is expected to induce a unique light signature in the detector, but dedicated searches have not yet succeeded. This work, for the first time, provides a tool for an all-embracing shower type classification of high energy cascades in IceCube.

Although resembling in topology, cascades from tau neutrino charge current, electron neutrino charge current and all flavor neutral current interactions differ in particle composition. This analysis uses the neutron content as a shower type indicator. Neutrons produced in neutrino interactions eventually capture on hydrogen nuclei in the ice, which give off a 2.2 MeV gamma quantum upon de-excitation. The photon is emitted $\mathcal{O}(100\mu s)$ delayed with respect to the cascade's prompt light signal, inducing a detectable afterglow - the so-called *neutron echo*. Since standard readout windows in IceCube are only open for around $10\,\mu s$, the automated request and processing of one second of unfiltered raw data, so-called *HitSpool* data, for high energy events has been implemented and monitored within the scope of this thesis to record the delayed light signal.

In this work, statistical analyses are used to determine the shower type composition of a high energy cascade sample and to probe two dark matter scenarios, in which Ice-Cube can be understood as a giant direct detection experiment for *boosted dark matter* particles. In Chapter 1 and Chapter 2, the thesis starts with a summary of the physics of neutrinos and their high energy interactions as well as the different neutrino-induced shower types. The IceCube detector, its data acquisition systems, neutrino sources and neutrino event signatures are presented in Chapter 3. To collect the current knowledge on the astrophysical neutrinos flux, Chapter 4 provides an overview of the results from dedicated IceCube analyses and furthermore introduces IceCube's *High Energy Starting Event* (HESE) selection, which facilitated the discovery of astrophysical neutrinos and which is also employed in this work. The theoretical background on neutron capture processes at IceCube conditions and their delayed light signal can be found in Chapter 5. The chapter additionally discusses alternative sources of delayed photons and explains how IceCube's HitSpool data acquisition stream is used for the recording of the afterglow of high energy events identified by the so-called HESE filter. Chapter 6 describes the stand-alone echo simulation as well as its integration into the full IceCube detector simulation chain. The tools needed for the final shower type analysis are provided in Chapter 7. This chapter first introduces the fundamental approach and subsequently explains its implementation. A fast algorithm, developed to make the analysis computationally and temporally feasible, is presented. The chapter furthermore provides the statistical methods applied in the individual event and ensemble analysis and discusses the involved systematic uncertainties. The analysis procedure is demonstrated by means of a representative simulation dataset, which is additionally used to cross-validate the developed routine. The chapter concludes by presenting the probed boosted dark matter models in greater detail and discusses the statistical methods used to test them. The final data analysis is contained in Chapter 8. As a first step, an atmospheric muon dominated dataset with photon counts below the analysis' charge threshold is investigated. An excess of delayed light, which was observed in this dataset, is tested with regard to its consistency with the neutron echo hypothesis. At last, a dataset containing 13 high energy cascade events recorded between February 25, 2016 and December 31, 2017 is analyzed in detail using the methods discussed in the previous chapters. The consistency with the most common assumption for astrophysical neutrino production as well as boosted dark matter scenarios is probed by means of a likelihood and an alternative Markov Chain Monte Carlo approach. The work concludes with a summary, improvement opportunities and analysis prospects focusing on IceCube's planned high energy extension *IceCube-Gen2*.

1

Neutrinos and their interactions at high energies

Neutrinos are elusive particles with tiny masses[1]. They carry no electric or color charge and thus interact with matter only via the weak force. The consequently small interaction cross section on the order of $10^{-35}\,\mathrm{cm}^2$ lets neutrinos traverse matter with a very small chance of being deflected or absorbed. It took physicists more than 25 years to provide experimental proof for the particle that was postulated by W. Pauli in 1930. The neutrino's reluctance to interact can, however, be exploited and opened up a new field of experimental physics: Neutrino astronomy. Earth-bound neutrino telescopes provide a valuable probe of astrophysical phenomena such as supernovae or gamma-ray bursts. The detection and study of extraterrestrial neutrinos thus helps to settle many open questions on the nature of the astrophysical particle accelerators.

1.1 Neutrinos within the Standrad Model and neutrino oscillation

The Standard Model of particle physics, in short Standard Model (SM), organizes all known elementary particles in a structured framework as shown in Figure 1.1. Fermions and bosons, which carry half-integer and integer spin, respectively, build the two main particle groups. Neutrinos, as spin-1/2 elementary particles, belong to the Standard Model's fermion content. The fermion group itself is composed of a quark and lepton group, each subdivided into three generations. A leptonic generation pairs a charged lepton with a neutrino of the same *flavor*:

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \qquad .$$

The neutrino flavor eigenstates are eigenstates of the weak force and a superposition of three mass eigenstates ν_1, ν_2, ν_3. This quantum mechanical phenomenon allows neutrinos produced in a certain flavor eigenstate to *oscillate* into a different flavor eigenstate. The effect of neutrino oscillations was experimentally confirmed in 1998 [2] and, as a non-zero neutrino mass is a prerequisite for the phenomenon, indirectly proved that at least two neutrinos are massive. The Standard Model, in which neutrinos initially appeared massless, had to be adjusted accordingly. One question that has not been settled so far is whether neutrinos are their own antiparticles, i.e. are of Majorana nature.

---

[1]Individual neutrino masses have not been experimentally determined so far. The best limit on the sum of neutrino masses, obtained from an analysis of cosmological observables, is $\sum m_\nu < 0.23\,\mathrm{eV}$ [1].
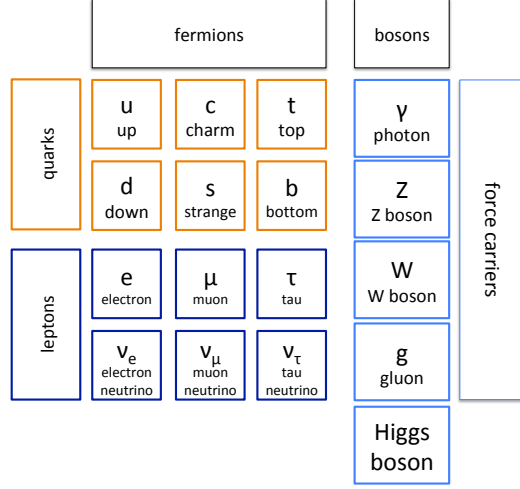
Figure 1.1: Overview of the Standard Model particle content. The original composition, developed in the 1970's, has been extended by the later discovered Higgs boson.

Studying the astrophysical neutrino signal, neutrino oscillation plays a special role as it changes the flavor composition of the neutrino flux on its journey from the production site to the Earth. The probability for a neutrino to oscillate from flavor eigenstate α to flavor eigenstate β along the flight length L is given by [3]:

$$
\begin{aligned}
P_{\nu_\alpha \to \nu_\beta}(L) = \quad & \delta_{\alpha\beta} \\
& - 4 \sum_{j>i} \Re(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}) \sin^2\left(\frac{\Delta m^2_{ij} L}{4 E_\nu}\right) \\
& + 2 \sum_{j>i} \Im(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}) \sin\left(\frac{\Delta m^2_{ij} L}{2 E_\nu}\right) \quad ,
\end{aligned}
\tag{1.1}
$$

where δ is the Dirac phase[2], U denotes the lepton mixing matrix[3], $\Delta m^2_{ij} = m^2_i - m^2_j$ is the difference of the squared neutrino masses and E_ν stands for the energy of the neutrino mass eigenstate. The sign of Δm^2_{ij} has only been determined for Δm^2_{12} by studying matter effects[4] in solar neutrino oscillation. The current global best-fit result

---

[2]If neutrino oscillation is CP-invariant, i.e. $\delta = 0$, the imaginary term in Equation 1.1 vanishes.

[3]commonly referred to as PMNS matrix, named after the physicists B. Pontecorvo, Z. Maki, M. Nakagawa and S. Sakata.

[4]In dense stellar objects like the Sun, $\overset{(-)}{\nu}_e$ experience a larger potential compared to the other flavor eigenstates, as they additionally interact with electrons present in matter. This leads to a larger effective mass for the electron flavor and thereby impacts the neutrino oscillation parameters. The phenomenon is named MSW effect, after the physicists S. Mikheyev, A. Smirnov and L. Wolfenstein.

yields $\Delta m_{21}^2 = +7.37 \times 10^{-5}\,\mathrm{eV}^2$ [4]. In the $\nu_\mu \to \nu_\tau$ oscillation in a low-density environment, which is the dominant transition in atmospheric neutrino oscillation, Δm_{31}^2 is accessible, however, only in terms of an absolute value. The current best-fit result states $|\Delta m^2| := |m_3^2 - (m_1^2 + m_2^2)/2| = 2.525\,\mathrm{meV}^2$ [4]. The so far undetermined sign of Δm_{31}^2 leaves two possibilities: either $m_3^2 > m_2^2 > m_1^2$ or $m_2^2 > m_1^2 > m_3^2$. The first scenario is referred to as *normal hierarchy* (NH), the second as *inverted hierarchy* (IH). Only in the case that the lowest mass eigenstate has a vanishing mass, the two remaining masses can be calculated individually. For the normal hierarchy and $m_1 = 0$, this yields $m_2 \approx 8.6\,\mathrm{meV}$ and $m_3 \approx 50\,\mathrm{meV}$. For $m_3 = 0$ in the inverted hierarchy, one finds that $m_1 \approx m_2 \approx 50\,\mathrm{meV}$ [5].

To calculate the flavor composition of the astrophysical neutrino flux detected at Earth, it is intuitive to consider neutrinos composed of wave packages representing different mass eigenstates. Due to the different masses, the wave packages have different group velocities, which lets their separation increase as they propagate through space. When a wave package no longer overlaps with others, its quantum mechanical interference with other mass eigenstates stops. Since the coherence length of the mass eigenstate packages is smaller than the size of the solar system [6], astrophysical neutrinos arrive at the Earth in a decoherent state, i.e. are no longer oscillating. For this case, the average of Equation 1.1 yields a constant transition probability of [7]

$$P_{\nu_\alpha \to \nu_\beta} = \sum_j |U_{\alpha j}|^2 |U_{\beta j}|^2 \qquad . \tag{1.2}$$

With recent best-fit values for the entries in the PMNS matrix, the astrophysical neutrino flux composition at the Earth's surface reads:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}_{\mathrm{NH}}^{\mathrm{Earth}} = \begin{pmatrix} 0.56 & 0.23 & 0.21 \\ 0.23 & 0.39 & 0.38 \\ 0.21 & 0.38 & 0.41 \end{pmatrix} \cdot \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}^{\mathrm{production\ site}}$$

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}_{\mathrm{IH}}^{\mathrm{Earth}} = \begin{pmatrix} 0.56 & 0.18 & 0.26 \\ 0.18 & 0.44 & 0.38 \\ 0.26 & 0.38 & 0.36 \end{pmatrix} \cdot \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}^{\mathrm{production\ site}} \qquad , \tag{1.3}$$

assuming normal and inverted hierarchy, respectively [8]. Different models for the astrophysical neutrino production and the resulting flavor composition at Earth are presented in Section 4.2.

1.2   Neutrino interactions
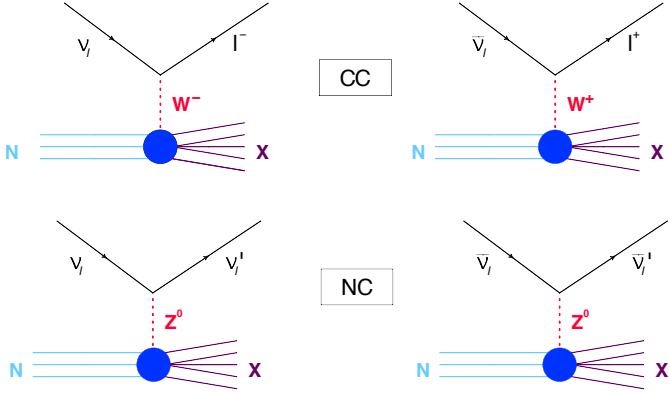
Figure 1.2: Illustration of a neutrino-nucleon (N) interaction. CC interactions convert the incoming neutrino into its charged leptonic partner, while NC interactions only impact the neutrino momentum. At all interaction vertices, a hadronic cascade (X) is initiated by the debris of the target nucleus.

In the Standard Model, neutrinos interact with matter via the exchange of the weak force mediators: the $W^{\pm}$ and Z boson. The corresponding interaction is named charged current (CC) interaction for the W boson exchange and neutral current (NC) interaction for the Z boson exchange (see Figure 1.2). In the energy regime relevant in this work, neutrinos are able to resolve the interior of nucleons in the target, thus interacting with their quarkonic constituents. The cross section for this *Deep Inelastic Scattering* (DIS) of a neutrino on an isoscalar nucleon in terms of the dimensionless Bjorken scaling variables $x = Q^2/(2M\nu)$ and $y = \nu/E_\nu$ reads [9]:

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

$$(1.4)$$

where E_ν is the energy of the incoming neutrino, $-Q^2$ is the Lorentz invariant momentum transfer, $\nu = E_\nu - E_l$ (CC) or $\nu = E_\nu - E_{\nu'}$ (NC) denotes the energy transferred to the target in its rest frame, M is the mass of the nucleon, $M_{W,Z}$ is the mass of the weak force mediators and G_F=1.16632 GeV^{-2} represents the Fermi constant. $q(x,Q^2)$, $\overline{q}(x,Q^2)$, $q^0(x,Q^2)$ and $\overline{q^0}(x,Q^2)$ are quark distribution functions in a proton, which depend on the momentum fraction x carried by the quark.

At the TeV energy scale, the cross section no longer increases linearly with energy due to the rising relevance of Q^2 in addition to the exchange boson mass in the propagator term in Equation 1.4. This leads to a suppression of the cross section for $E_\nu > 10$ TeV [10].

Beside DIS, electron antineutrinos can interact with atomic electrons in the target material. This is a subdominant channel for most of the energy range, but becomes a leading process at around 6.3 PeV, when the center of mass energy allows for the resonant production of a real W^- boson on ambient electrons in matter (see Figure 1.3).

This mechanism was first proposed by Sheldon L. Glashow in 1960 [11] and is thus referred to as the *Glashow resonance*[5].
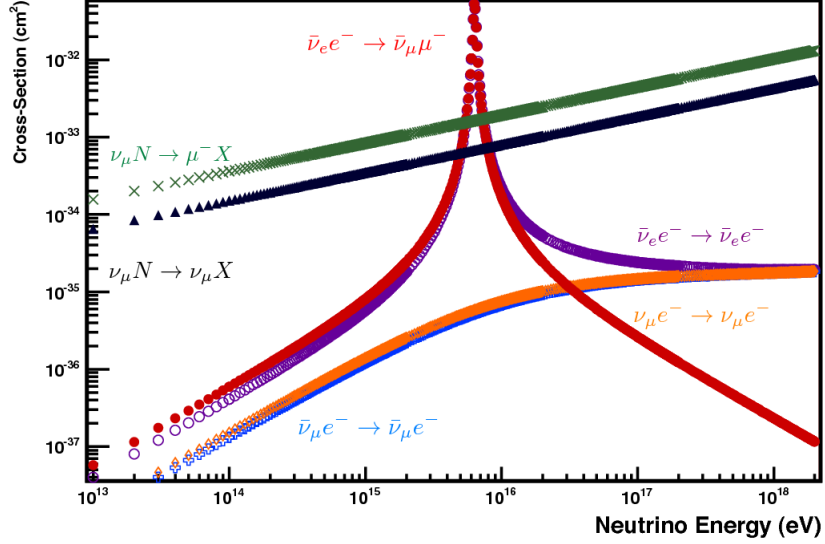
Figure 1.3: Neutrino scattering cross sections in the ultra high energy regime above $10\,\mathrm{TeV}$. At $6.2\,\mathrm{PeV}$, the resonant production of the W^- via neutrino-electron-scattering becomes the dominant interaction process. The graphic is taken from [10] with results from [12, 13].

1.3 Neutrino detection

Since neutrinos are only subject to the weak force, they cannot be observed directly. However, charged particles generated in a neutrino interaction produce a detectable light signal predominantly via Cherenkov radiation.

Cherenkov radiation occurs when charged particles traverse a dielectric medium with a velocity greater than the local phase velocity of light in that medium, c/n, with the medium's refractive index n. Electromagnetic waves emitted by atoms, upon de-excitation after polarization by the passing particle, superimpose coherently. Constructive waves form a light cone as illustrated in Figure 1.4. A charged particle moving at a velocity v from O_1 to O_5 in time t, creates a light cone with Cherenkov angle θ_C given

---

[5]The Glashow resonance at $s = M_\mathrm{W}^2$ extends over a resonance region of $(M_\mathrm{W} - 2\Gamma_\mathrm{W})^2/2m_\mathrm{e} = 5.7\,\mathrm{PeV} \lesssim E_\nu \lesssim (M_\mathrm{W} + 2\Gamma_\mathrm{W})^2/2m_\mathrm{e} = 7.0\,\mathrm{PeV}$ and has a peak cross section of approximately $5 \times 10^{-32}\,\mathrm{cm}^2$.

by

$$\cos(\theta_C) = \frac{c/n \cdot t}{O_1 O_5} = \frac{c/n \cdot t}{v \cdot t} = \frac{1}{\beta n} \qquad . \tag{1.5}$$
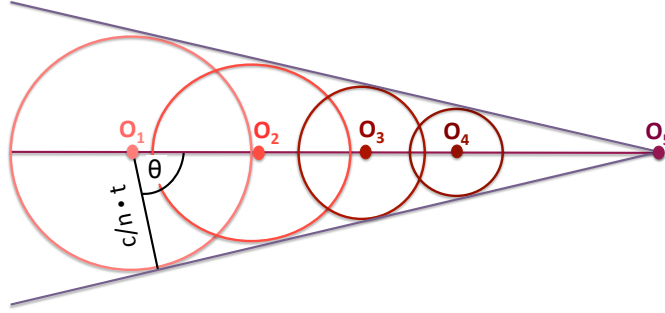
Figure 1.4: Applying Huygens principle to the electromagnetic waves emitted by de-exciting atoms shows the formation of a light cone if the source of excitation (i.e. a charged particle) travels at a velocity greater than the speed of light in the medium with refractive index n.

From this formula it can be deduced that no Cherenkove radiation is observable for $v < c/n$. This translates to a kinetic energy threshold of $E_{th,\mu} \approx 55\,\mathrm{MeV}$ and $E_{th,\mathrm{e}} \approx 0.26\,\mathrm{MeV}$ for muons and electrons, respectively, in water and ice [14]. The spectrum of Cherenkov photons N per wavelength interval $d\lambda$ and unit path length dx from a traversing particle with charge $z \cdot e$ is described by the Frank-Tamm formula [15, 16]:

$$\frac{d^2 N}{dx d\lambda} = \frac{2\pi \alpha z^2}{\lambda^2} \left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right) \qquad . \tag{1.6}$$

Here, the refractive index n is wavelength dependent and α denotes the fine structure constant. In IceCube's most sensitive wavelength range from around $300\,\mathrm{nm}$ to $500\,\mathrm{nm}$, a relativistic muon induces $250\,\mathrm{cm}^{-1}$ Cherenkov photons assuming a constant refractive index of $n \approx 1.33$ for ice [14]. It should be noted that Cherenkov radiation plays a negligible role with regard to leptonic energy losses [16] (see also Section 2.1.1).

2

Electromagnetic and hadronic showers

Particle showers following a neutrino interaction can be divided into two fundamental categories: Electromagnetic and hadronic showers. In IceCube, both shower types are referred to as *cascade*. Electromagnetic cascades are induced by an electron (positron) generated in electron neutrino (electron antineutrino) CC interactions. Hadronic cascades, in contrast, originate from the fragmentation of the nucleus in every neutrino interaction. The cascade types are, however, not *pure*, as every hadronic shower has an electromagnetic part from the internal production of electrons and photons, and photonuclear reactions are a source of a hadronic component in electromagnetic showers. This sections collects theoretical background information on both shower types and presents the main cascade-inducing reactions in IceCube.

2.1 Electromagnetic cascades from ν_e CC interactions

The high energy electron or positron produced in an electron flavor charged current interaction will set off an electromagnetic shower in the ice by emitting bremsstrahlung in the electromagnetic field of atoms. The so created photons will subsequently produce an $e^{\pm}$ pair via pair production, inducing a new shower level containing twice as many particles as the previous one. The rather regular cascading of electrons, positrons and photons defines an electromagnetic shower. The development of the shower ends once the particle energies are below a critical energy.

2.1.1 Passage of electrons, positrons and photons through matter

Although Figure 2.1 shows the fractional energy loss of electrons and positrons traversing lead, it is suitable to summarize possible mechanisms and illustrates their energy dependence: at low energies, $e^{\pm}$ predominantly interact with atomic electrons in the ice. Although scattering (Møller scattering for electrons, Bhabha scattering for positrons) and positron annihilation also contribute, ionization of atoms in the ice is the major energy loss process for low energy $e^{\pm}$. Above the *critical energy E_C*, Bremsstrahlung takes over as the primary energy loss process. E_C is defined as the energy at which the rate of collision losses through ionization and excitation equals $X_0 E$, with the electron energy E and the radiation length X_0. X_0 is defined as the mean distance over which a high-energy electron loses all but $1/e$ of its energy by Bremsstrahlung, and $7/9$ of the mean free path for pair production for high energy photons [17]. In ice, these characteristic quantities are $E_{C,e^-} = 78.60\,\text{MeV}$, $E_{C,e^+} = 76.51\,\text{MeV}$ and $X_0 = 39.3\,\text{cm}$ [18].
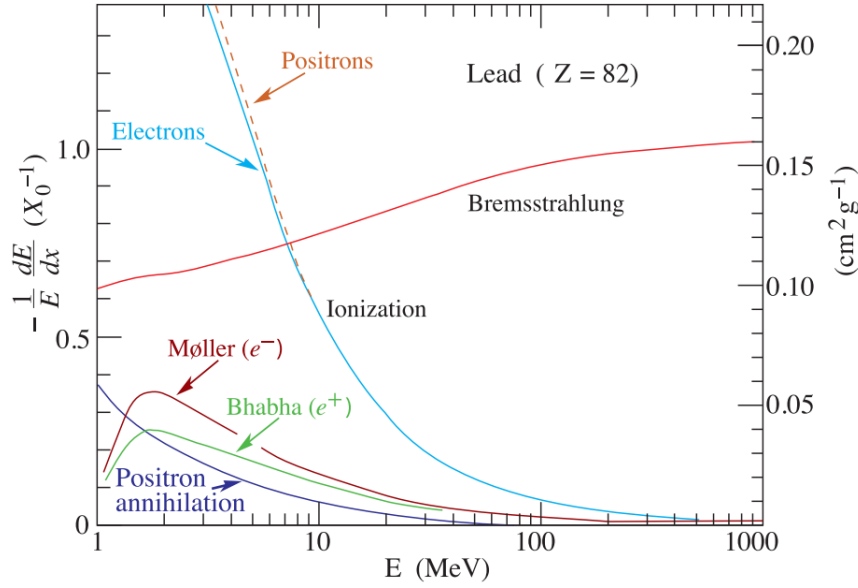
Figure 2.1: Fractional energy loss per radiation length in lead as a function of electron or positron energy [16].

Low energy photons traversing matter lose their energy mainly through the photoelectric effect and furthermore experience Compton and Rayleigh scattering as well as photonuclear absorption. With increasing photon energy, these processes become subdominant compared to pair production in the surrounding fields of atoms.

At the very high energy end, above roughly 10^{16} eV in ice [19], the Landau-Pomeranchuk-Migdal (LPM) effect[1] suppresses pair production as well as bremsstrahlung processes. From 10^{20} eV on, the photonuclear and electronuclear cross section becomes larger than that for pair production and bremsstrahlung, respectively, and electrons additionally lose energy via direct pair production $e^- N \rightarrow e^+ e^- e^- N$, with N denoting a nucleus [19]. A further hadronic contribution appears when photons reach even higher energies allowing the fluctuation into quark-antiquark pairs which subsequently form real ρ^0 mesons.

---

[1]The LPM effect refers to the quantum mechanical interference of a particle's scattering steps, which can thus no longer be described as independent events.

2.1.2 Electromagnetic cascade development

The total track length of an electromagnetic shower increases almost linearly with the primary energy E_0 and can be parametrized by:

$$\hat{l}(E_0) = \alpha \cdot E_0^{\beta} \qquad , \qquad (2.1)$$

with β being consistent with unity to a level of 10^{-5} for different primary particles (e$^-$, e$^+$, γ) [20]. $\hat{l}$ denotes the charged-particle track length corrected by the Franck-Tamm factor, which accounts for the variation of the Cherenkov light yield due to different particle velocities. The parameter α is also constant up to a level of 10^{-3} and was determined to be 532.1 ± 10^{-3} cm/GeV [20].

2.2 Hadronic cascades from NC interactions

A neutrino interacting with a nucleus in the ice initiates a hadronic cascade by fragmenting its interaction partner. Hadrons, which will form from the spallation products, will subsequently interact with surrounding nuclei, thereby creating even more particles. This chain leads to the development of a particle shower in analogy to the one induced by electrons. However, since a variety of particles is produced in hadronic showers and many processes can contribute, the hadronic shower development is complex and large fluctuations result in a more *frayed* shower profile compared to electromagnetic cascades (see Figure 2.2).

Because the nuclear interaction length (90.77 cm in ice [18]) is larger than the electromagnetic radiation length, hadronic cascades are longer than electromagnetic showers at low energies. Since the LPM effect does not significantly impact hadronic showers, they do not experience the elongation effect of electromagnetic cascades and do not extend over more than 40 m in ice [19]. The energy at which the electromagnetic exceeds the hadronic shower length is about 10^{19} eV [19].

Just as electromagnetic cascades have a hadronic component, hadronic showers are not exclusively hadronic. The prompt decay of neutral π^0 mesons into two photons triggers the development of an intrinsic electromagnetic cascade. This electromagnetic component increases with cascade energy due to the rising number of generated π^0. The energy dependence of the interplay of electromagnetic and hadronic processes and the resulting variation in the Cherenkov light yield can be described via the ratio of the total hadronic over the total electromagnetic track length:

$$F = \frac{T_{had}}{T_{em}} \qquad . \qquad (2.2)$$

This can be rewritten as

$$F = F_{em} + F_{had} \cdot f_0 = F_{em} + (1 - F_{em}) \cdot f_0 \qquad , \qquad (2.3)$$
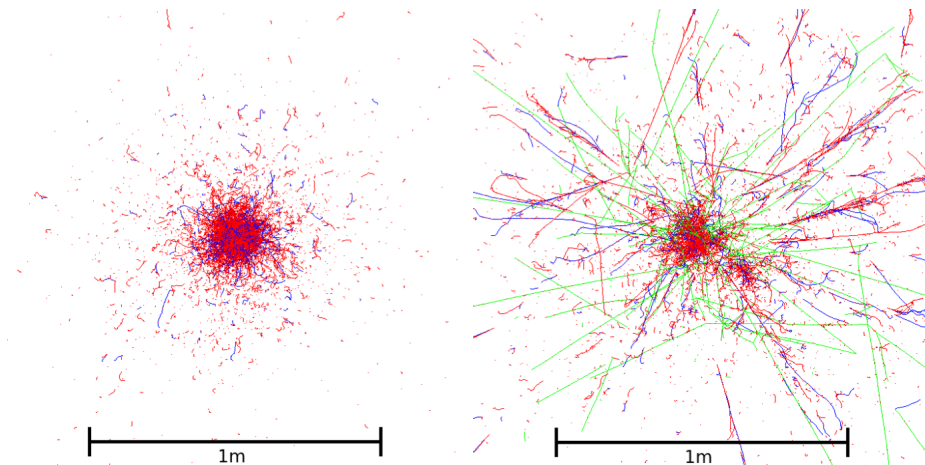
Figure 2.2: The transversal view on an electromagnetic (left) and a hadronic (right) cascade illustrates the more *frayed* and wider shower profile of the hadronic shower. The cascades shown were simulated with Geant4 using a primary particle of 100 GeV energy. Electrons and positrons are displayed in red and blue, respectively, and tracks of charged hadrons are colored in green [20].

where F_{em} and F_{had} are the energy fraction going into the electromagnetic and hadronic part of the cascade, respectively, and f_0 is the relative Cherenkov light yield of the purely hadronic component. When parametrizing F_{em} as

$$F_{em} = 1 - \left(\frac{E_0}{E_S}\right)^{-m} \qquad , \qquad (2.4)$$

with the initial energy E_0, the phenomenological parameters f_0, m and E_S can be obtained from simulation results. Values for different primary particles can be found in [20], where these parameters were fitted to Geant4 simulations. The energy dependence of the Cherenkov light yield of hadronic cascades relative to electromagnetic showers is shown in Figure 2.3.

2.3 Further cascade-inducing processes

Besides the two main shower inducing mechanisms presented in the previous sections, decaying τ leptons generated in ν_τ CC interactions as well as high energy muons may also initiate cascades.

2.3.1 Cascades from ν_τ CC interactions

A two-component light pattern uniquely identifying the τ lepton created in a ν_τ CC interaction (see Section 3.4) is only expected above a few hundred TeV deposited energy,
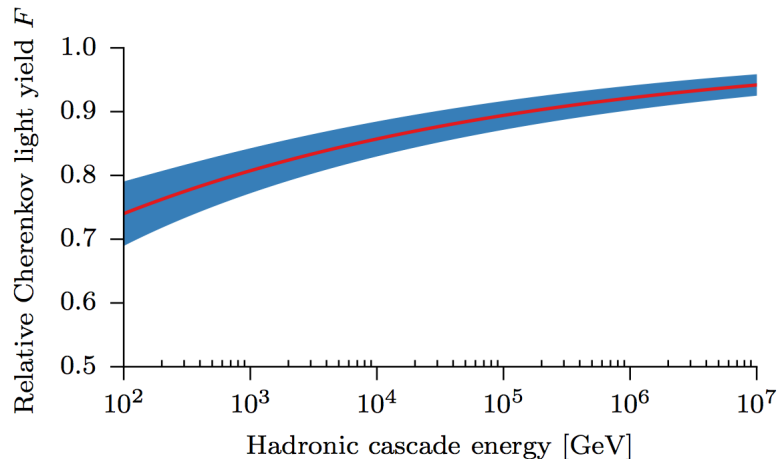
Figure 2.3: The visible fraction of Cherekov light in hadronic cascades relative to electromagnetic showers of the same energy [21]. Parameters are obtained from a fit to Geant4 simulations of hadronic jets following a neutrino-nucleon interaction simulated with PYTHIA 6.4. [22]. The red line shows the mean of F and the blue band its standard deviation. Variations in F are mostly due to fluctuations in the number of π^0 particles produced in hadronic showers.

where the τ decay length is larger than $20\,\text{m}$. No such events were found in dedicated IceCube searches so far (see e.g. [23, 24]). Below this energy scale, the cascade from the neutrino interaction, the track of the created τ lepton and the second shower from its decay can not be resolved by the detector and appear as a single cascade. With a branching fraction of around $17\,\%$, the τ can also decay into a muon, which subsequently induces a track-like pattern indistinguishable from a ν_μ CC interaction. In summary, the dominant τ decay modes together with their branching fraction and observable event signature are:

τ decay mode	branching fraction	signature in IceCube
$\mu^\pm + \overset{(-)}{\nu}_\mu + \overset{(-)}{\nu}_\tau$	(17.39 ± 0.04)	track
$e^\pm + \overset{(-)}{\nu}_e + \overset{(-)}{\nu}_\tau$	(17.82 ± 0.04)	electromagnetic cascade
hadrons $+ \overset{(-)}{\nu}_\tau$	$\approx 64\%$	hadronic cascade

Table 2.1: τ decay modes [16].

A combination of the above event signatures and the hadronic cascade at the neutrino interaction vertex will be observed in the detector.

13

Since the τ decays exclusively through the weak interaction, a tau neutrino or antineutrino is required in the final state to conserve lepton number. On average, around 25 % of the energy from the τ decay cannot be detected as it is carried by outgoing neutrinos [25].

2.3.2 Cascades associated with muons

At low energies, muons lose their energy mainly via ionization (Figure 2.4 (a)), which can be described as a continuous process. At energies typical for atmospheric muons, however, radiative processes (Figure 2.4 (b)-(d)) become the dominant energy loss mechanisms. For muons in ice, radiative processes dominate over ionization from 844 GeV muon energy on [26] and already contribute 95 % to the total losses in water at 20 TeV muon energy [27]. In contrast to the continuous energy loss through ionization, radiative processes have a stochastic nature and are associated with *catastrophic* energy losses inducing electromagnetic as well as - in the case of photonuclear reactions - hadronic showers. Muon induced hadronic cascades are of special importance in this analysis, since they mimic a neutrino NC interaction in background studies (see Section 8.1). Figure 2.5 displays the rate of radiative energy losses from atmospheric muons in Ice-Cube versus the deposited energy and emphasizes the potential of this type of event as being falsely identified as a cascade-like neutrino interaction. Photonuclear muon interactions contribute around 25 %, 20 % and 15 % to stochastic energy losses of 1 TeV, 10 TeV and 100 TeV, respectively.
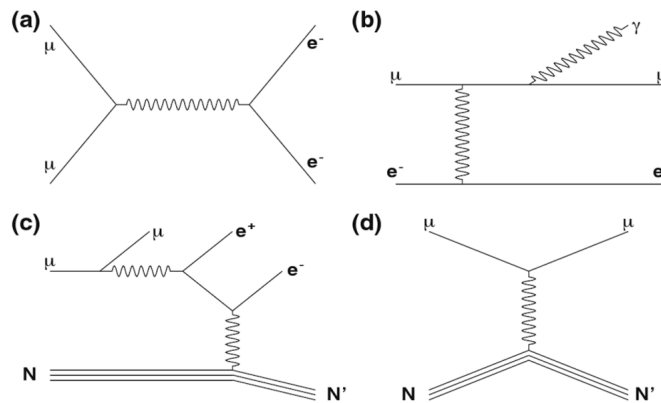
Figure 2.4: Feynman diagrams for muon energy loss processes in matter [28]: (a) muon scattering on electrons bound in atoms, (b) bremsstrahlung, (c) electron pair production and (d) photonuclear interaction. Here, N denotes a nucleon.
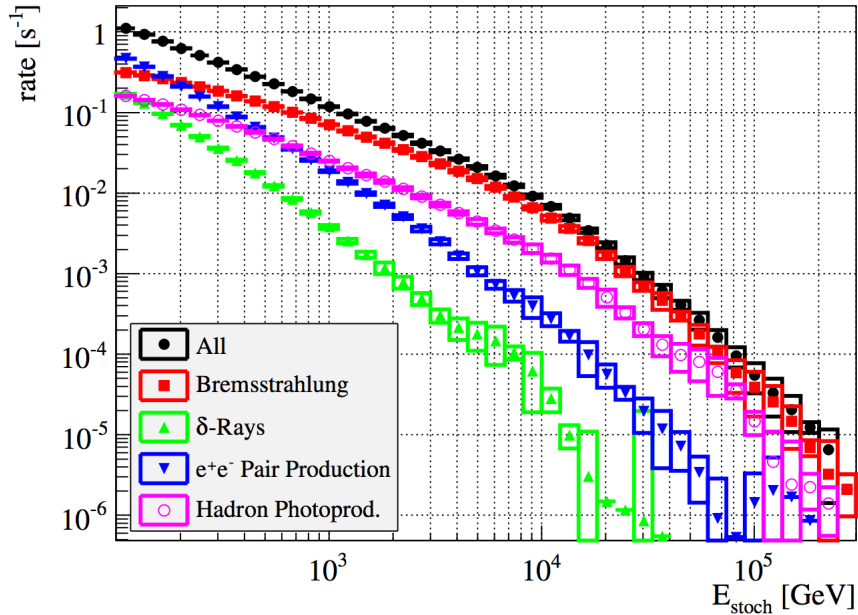
Figure 2.5: The energy spectrum of stochastic energy losses from simulated muon events with more than 1000 detected PE in IceCube [29]. δ-Rays refer to free atomic electrons generated via muon scattering (Figure 2.4 (a)), which received enough energy to distance themselves from the atom and undergo further interactions, thereby producing a detectable light signal.

2.3.3 Glashow resonance events

The signature of a Glashow resonance event (see Section 1.2) is determined by the decay channels of the W^- generated in the process $\bar{\nu}_e + e^- \to W^-$. The relevant branching ratios are:

W^- decay mode	branching fraction	signature in IceCube
$\tau^- + \bar{\nu}_\tau$	$(11.38 \pm 0.21)\%$	see Table 2.1
$\mu^- + \bar{\nu}_\mu$	$(10.63 \pm 0.15)\%$	track
$e^- + \bar{\nu}_e$	$(10.71 \pm 0.16)\%$	electromagnetic cascade
hadrons	$(67.41 \pm 0.27)\%$	hadronic cascade

Table 2.2: W^- decay modes [16].

With a nearly $70\,\%$ chance for a hadronic W^- decay, the detection of a roughly $6\,\mathrm{PeV}$ hadronic shower would strongly support the first detection of a neutrino interacting via this mode.

3

The IceCube neutrino telescope

IceCube is the world's largest neutrino telescope, using around $1\,\mathrm{km}^3$ of clear Antarctic ice as its detector material to catch elusive neutrinos or exotic particles. Its design was optimized to observe high energy neutrinos starting at the TeV scale and explore their yet unknown astrophysical origin. Apart of the study of astrophysical particle accelerators, IceCube's observable energy range extends to low enough values to investigate, for instance, general neutrino properties by analyzing the atmospheric neutrino spectrum.

3.1 The IceCube detector

IceCube [30–32] records Cherenkov radiation from neutrino interactions by means of 5160 *Digital Optical Modules* (DOMs) frozen into Antarctic ice in the immediate vicinity of the geographic South Pole. In its final configuration completed in 2011 (Figure 3.1), IceCube is composed of 86 cables - so-called *strings* - each holding 60 DOMs at a depth between 1450 m and 2450 m. Quad cables provide the DOMs' electric power as well as an inter-DOM and surface communication system.

Beside the 78 standard strings deployed nearly equispaced at a distance of 125 m from one another, IceCube contains the low energy infill *DeepCore*, with a reduced inter-string spacing of 41 m to 88 m. In DeepCore, the inter-DOM spacing of 7 m is substantially lower than the roughly 17 m on standard strings. Most of the DOMs in DeepCore are equipped with photomultipliers that possess an around 35 % higher quantum efficiency (QE).

On the surface of the ice sheet, located at the string positions, is the air shower detector *IceTop*. It serves also as a background veto and calibration tool and is composed of 81 stations each containing two tanks with two standard DOMs frozen into clear ice.
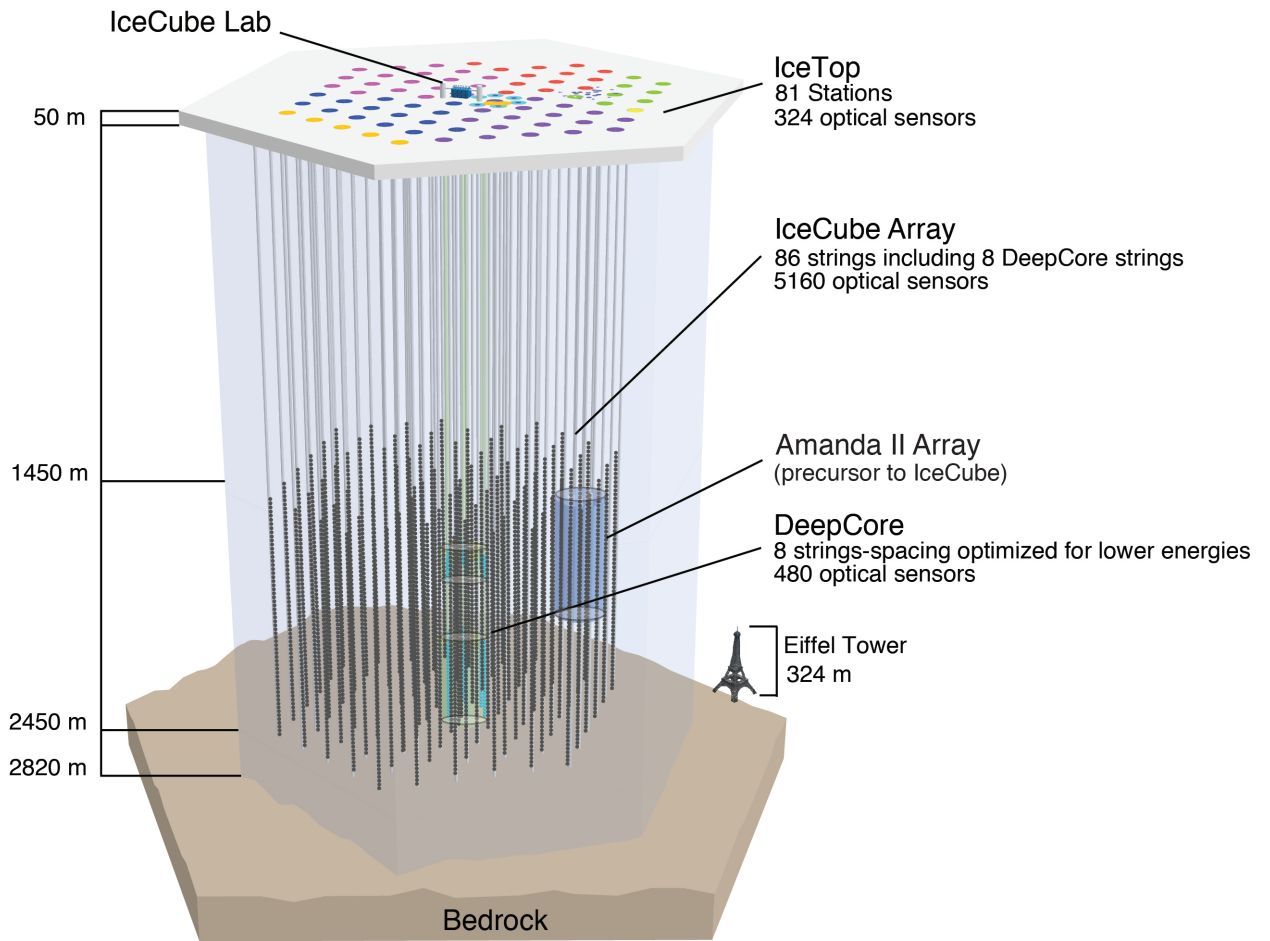
Figure 3.1: The IceCube detector. The IceCube Laboratory (ICL) at the surface houses IceCube's data acquisition system, where the signals of all DOMs are collected, buffered and processed.

3.2 Data acquisition

IceCube's standard optical module, as schematically shown in Figure 3.2, includes a 25 cm diameter Hamamatsu R7081-02 photomultiplier tube (PMT) within a 13 mm thick borosilicate glass vessel for the detection of photons. RTV (Room Temperature Vulcanization) silicone gel optically couples the PMT to its glass housing and serves as a shock absorber. The PMT receives its power from the high voltage (HV) generator and divider circuit and is shielded against the impact of the Earth's magnetic field by a mu-metal grid. The LED Flasher board is incorporated to generate controlled light pulses for calibration and ice properties studies. After setup, the DOM is filled with dry nitrogen

regulating its interior pressure to about 1/2 atmosphere to ensure mechanical stability.
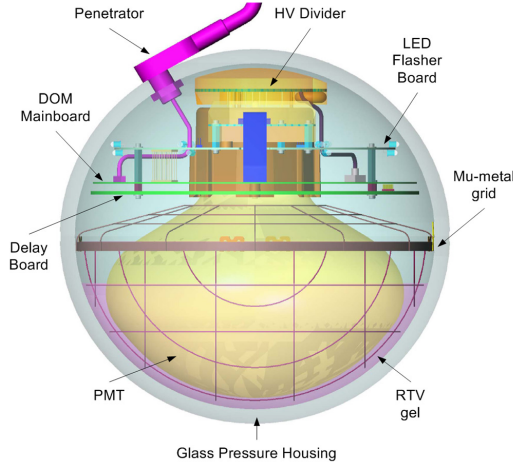
Figure 3.2: The IceCube DOM with its main components [30].

The main task of the assembled DOM is the recording, timestamping and extraction of waveform information of signals crossing the PMT's discriminator threshold[1] - so-called *hits* or *launches*. The analog PMT signal is split into three paths and passed to the DOM mainboard (MB) (see Figure 3.3). The upper path serves the triggering if the signal passes the PMT discriminator threshold. The digitization of the analog signal is performed on the remaining signal paths which enter two distinct units:

Analog Transient Waveform Digitizer (ATWD): The ATWD records the signal waveform in case the PMT threshold was exceeded. In order to provide enough time for this unit to receive the corresponding initiation from the trigger path, its input is delayed by 75 ns in the DOM delay board. After triggering, the ATWD collects the analog signal in 128 sampling capacitors at a rate of 300 MSPS (MegaSamples Per Second). This translates to the recording of a 427 ns long waveform in 3.3 ns wide bins. The waveform capture is performed in three channels at different amplification levels ($\times 0.25$, $\times 2$, $\times 16$), while the subsequent digitization by internal ADCs (Analog to Digital Converters) is run on the most amplified channel first and is only performed on lower amplification channels in case of a possible overflow for large signals. The fourth ATWD channel can receive eight different signals through multiplexers and is mostly used for calibration purposes.

Fast Analog to Digital Converter (FADC): In order to be able to capture signals lasting longer than can be covered by the ATWDs, the DOM MB additionally

---

[1]The configurable discriminator level is set to roughly 0.25 of the pulse height induced by a single photoelecton in the PMT.

includes an FADC. It samples at a lower rate of 40 MSPS for an, in general, arbitrarily long time span. However, the capture window is set to 256 samples, thus covering a time window of $6.4\,\mu$s in 25 ns bins. The input signal is broadened by a 180 ns pulse shaper to facilitate an arrival time estimation for isolated PMT output signals despite the FADC's coarse resolution [33]. The readout of the FADC is furthermore aligned with the ATWD readout.
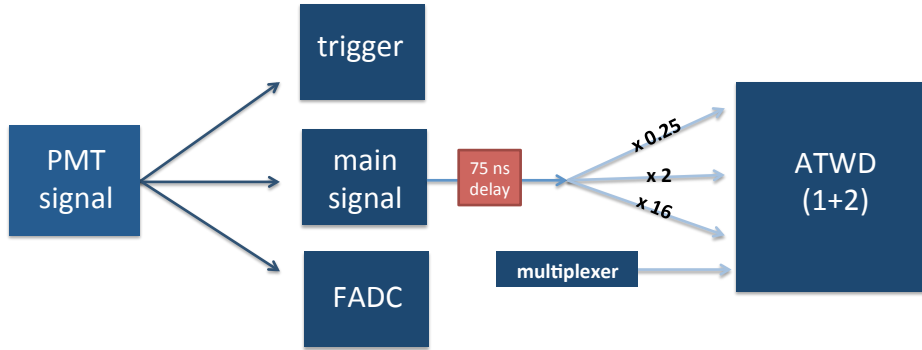
Figure 3.3: Schematic diagram of the processing flow of the PMT signal. The readout of the ATWDs as well as the FADC is controlled by an FPGA (Field-Programmable Gate Array) not shown in this sketch.

Since multiple photons can contribute to one hit, the digitized waveform may exhibit multiple *pulses* (see Figure 3.4). Due to the limited time resolution, a specific pulse can itself be formed by multiple photons, yielding a larger pulse height. The extraction of pulses [34] allows IceCube to calculate photon arrival times and charges. Waveform unfolding is run at the Pole and offline in Madison, Wisconsin.

19
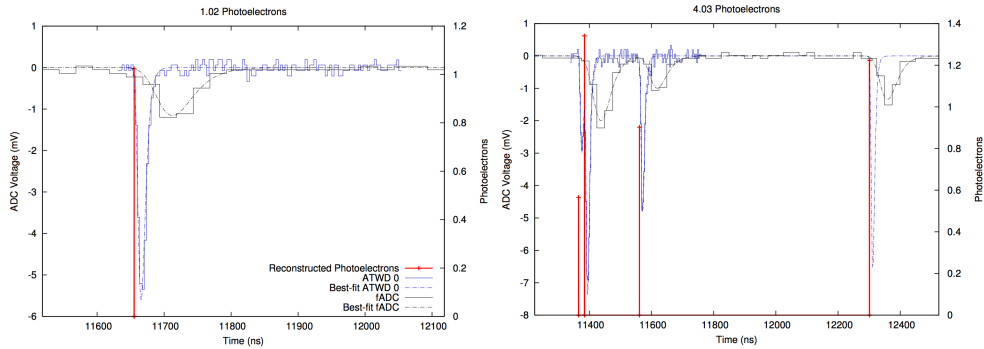
Figure 3.4: Examples of pulse extraction via waveform unfolding for a simple (left) and complex (right) waveform [34]. Solid and dashed lines show the digitized waveforms and their predicted shape given the reconstructed PMT hits, respectively. Red solid lines mark the reconstructed time and charge (right-hand axis) of the extracted pulses.

3.2.1 Detection limitations and spurious hits

Ideally, IceCube's DOMs would report every photon from an external light source that produces a photoelectron in the PMT. In practice, however, it is sometimes not possible to reconstruct the complete signal. Additionally, spurious hits are registered even if the surrounding is completely dark. The sources of these detection artifacts are discussed in the following paragraphs.

Saturation

Most neutrino interactions do not lead to more than a few hundred photons hitting single PMTs. For these input signals, the PMT output is proportional to the number of detected photons. However, very bright events can bring the PMTs experiencing the highest light incidence to their saturation region, where the output signal is no longer proportional to the number of incident Cherenkov photons. Figure 3.5 shows the saturation effect for IceCube PMTs as measured in the laboratory.

Late pulses and afterpulses

Electrons not passing in sequence through the amplification chain of the PMT's dynodes give rise to delayed signals. Photoelectrons scattering back from the first dynode before hitting it again experience a time delay which causes a delayed PMT output signal called *late pulse*. Around 4% of all hits can be attributed to this process, with the main part occurring 26-65 ns after the primary signal, a second peak at around 71 ns and a third contribution with 81-160 ns delay [35].

Afterpulses arise when electrons, accelerated between the dynodes, interact with residual gas molecules in the PMT. Ions produced in these interaction accelerate back towards the photocathode and release delayed electrons from it. The afterpulse spectrum can

20
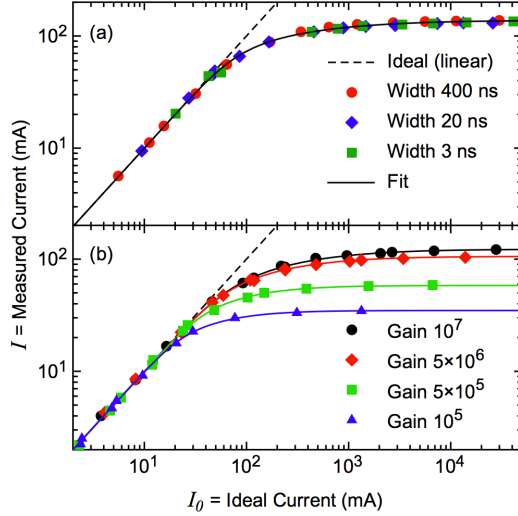
Figure 3.5: The saturation behavior has been studied by flashing IceCube PMTs with an LED [35]. The saturation effect is clearly visible when plotting the measured photocurrent versus the current expected from an ideal, linearly responding PMT. The dependence on different LED pulse shapes (a) as well as different PMT gains (b) was investigated.

be described by a series of Gaussian distributions modeling the afterpulses of different ion types [36]. While afterpulses are observed over a time span of $300\,\text{ns} - 11\,\mu\text{s}$ after the main signal, the most prominent peaks are located at $600\,\text{ns}$, $2\,\mu\text{s}$ and $8\,\mu\text{s}$. For less than around 10^6 primary photoelectrons, 0.06 single-photoelectrons per primary photoelectron are expected from afterpulses [35]. Rather than incorporating individual afterpulse spectra for the deployed PMTs, an average model is used for all DOMs in IceCube data simulation so far.

Deadtime

A DOM not being available for photon recording, leading to a so-called deadtime, can be due to two reasons: either the DOM electronics themselves are still busy processing the signal or no ATWD is available when being engaged in its time consuming hit processing. Such deadtime effects are negligible for the typical duration of IceCube events of $\mathcal{O}(10\,\mu\text{s})$. However, they play an important role when the time frame is significantly expanded.

The time span for which a DOM is unavailable after a launch depends on the *Local Coincidence* (LC) condition of the previous hit: for each hit, the electronics looks for an other hit in one of the two neighboring DOMs above and below the module within a time window of $\pm 1\,\mu\text{s}$. If such a hit is found, all hits fulfilling this criterion are labeled *Hard Local Coincidence* (HLC). Hits which do not meet the criterion are called *Soft*

Local Coincidence (SLC) hits. Since HLC hits are more likely to originate from physics events than from noise, the full waveform information (ATWD and FADC readout) is stored. If the HLC condition is not met within the coincidence window, the processing of the waveform is aborted to reduce the data load. Only the highest of the first 16 FADC samples and its two direct neighbors, from which a coarse charge information is calculated, are stored for these events [33].

The contributions and calculation of the respective deadtimes are summarized in Table 3.1. The various denotations used therein are:

Ad: time needed to digitize one ATWD channel (29 μs)

Ar: time needed to restart an ATWD (225 ns)

Ac: time needed to clear one ATWD channel (950 ns)

Arec: ATWD recording time (427 ns)

CLr: cable length correction, which accounts for the time needed to transfer the LC signal to the DOM's furthest neighbor (1325 ns)

Pw: post time window for the LC decision (1 μs)

Cl: clock cycle length of the FPGA (25 ns)

	DOM busy for	ATWD busy for
after SLC hit	CLr + Pw + (5+2)·Cl 2.5 μs	CLr + 5 Cl + (4-nD)·Ac + Ar 5.475 μs
after HLC hit	256 FADC samples + 2·Cl 6.45 μs	nD · Ad + Ar + (4-nD)·Ac + Arec 32.502 μs (nD=1) 116.652 μs (nD=4)

Table 3.1: For how long the DOM itself or an ATWD is engaged in processing after an HLC and SLC hit is summarized in the above table. Blue numbers below the formulae state the corresponding time period when no ATWD channel overflows and the number of ATWD channels to be digitized (nD) is one. If all four ATWD channels have to be read out (nD=4), the ATWD is busy for around 116 μs (red number). In the SLC case, nD is alway zero.

Due to the long time required for the digitization of the ATWD signal, a second ATWD is in place reducing the period in which no photons can be recorded to 19.7 μs in the scenario of three subsequent HLC hits and one read-out ATWD channel (see Figure 3.6). When, however, the lowest amplification channel has to be digitized for very bright events, the deadtime in the case of three subsequent HLC hits is as large as 104 μs.
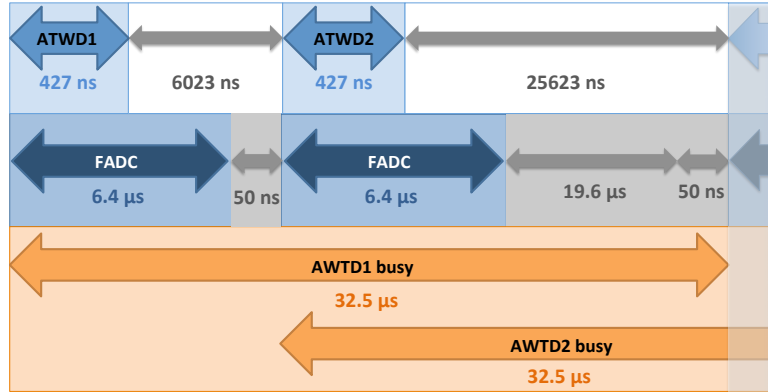
Figure 3.6: In the case of a second HLC hit, an additional ATWD is in place to capture the waveform. If only one ATWD channel is read out, i.e. no overflow occurred in the most amplified channel, ATWD1 is available only after processing is complete. The total deadtime (gray bands) until the third HLC hit is recorded thus adds up to $19.7\,\mu$s.

Dark noise

A PMT output signal in the absence of an external photon source is referred to as *dark noise.* Among others, dark noise can be caused by thermionic emission of electrons, noise from the DOM's electronic devices and light producing effects such as luminescence and Cherenkov emission powered by radioactive decays, e.g. of Potassium (^{40}K), in the glass and pressure sphere [32, 37]. For standard IceCube DOMs, the dark count rate is approximately $560\,$Hz, while high QE DOMs deployed in DeepCore show a noise rate of around $780\,$Hz. In data, a correlated noise contribution is observed besides the Poissonian component (see Figure 3.7). While the exact origin is yet undetermined, the most favored explanation sees correlated noise as a consequence of luminescence in the glass subsequent to a radioactive decay, e.g. in the Uranium or Thorium chains [32, 37]. Correlated noise occurs at a rate of $280 - 340\,$Hz in addition to the Poissonian contribution at $230 - 250\,$Hz.

3.2.2 HitSpooling

HitSpooling is a data acquisition operation originally designed to catch the overall rise in noise rate caused by the low energy neutrino flux from supernova events [38]. It allows one to read out the untriggered, raw data of the entire detector in a selectable time
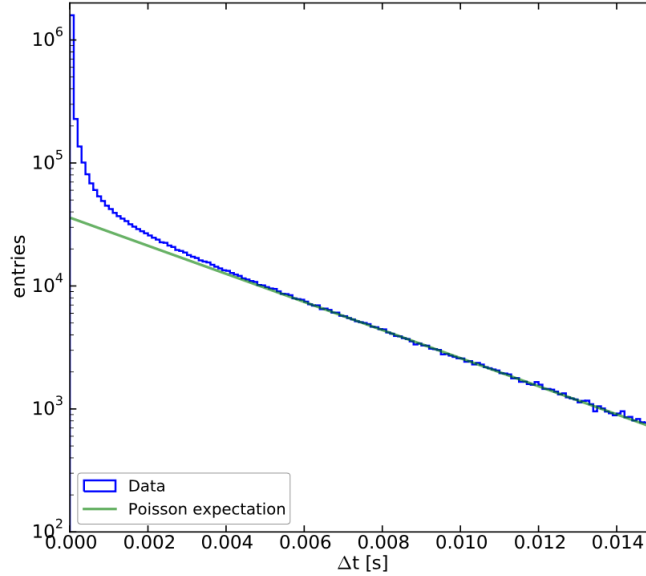
Figure 3.7: The distribution of time intervals between consecutive noise hits demonstrates the presence of correlated noise as an excess over the Poissonian expectation (green line) at small time intervals [32].

interval. Besides data taking by IceCube's standard data acquisition system (DAQ), hits are constantly *spooled*[2] to disks on the *DOMHubs* - computers located in the ICL, which serve as a communication interface to the DOMs on their assigned string. On each *DOMHub* runs a software component named *StringHub* composed of the frond end *Omicron* and the back end *Sender* [32]. *Omicron* manages the DOMs on its string and passes their hits to the *Sender* in a chronological stream, which subsequently buffers the hits and prepares HLC hits for local trigger application. Between *Omicron* and *Sender*, hits are written to the HitSpool buffer (see Figure 3.8). In order to initiate the read out of the spooled data, a daemon, which receives the request and hands it over to the computer, is running on each *DOMHub*.

When a selected time interval of HitSpool data is collected subsequent to a read-out request, the raw data get compressed and moved to a directory, in which the *JADE* (JAVA Archival and Data Exchange) software picks up files for satellite transfer to the North. *JADE* and its precursor *SPADE* (SouthPole Archival and Data Exchange) manage the data at the Pole by organizing its satellite or real-time transmission or local archiving.

As the Hitspool system runs independent from IceCube's DAQ, the risk of undesired interference between the two systems is low and HitSpool data can serve as a back-up in

---

[2]After a time period of currently up to 150 hours, data in the ring buffer will be overwritten.

the worst case scenario of a DAQ blackout. More details on the usage of the HitSpool system in high energy events is given in Section 5.2.
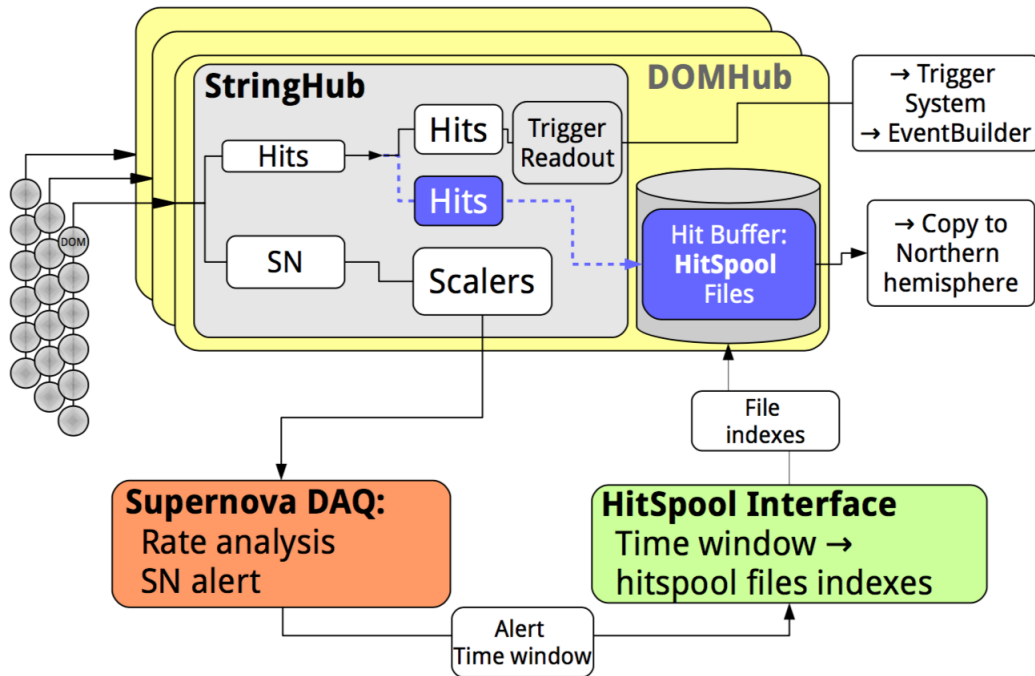
Figure 3.8: The diagram illustrates how HitSpool data requests by the supernova (SN) DAQ are implemented on the *DOMHub* [38]. In addition to SN alerts, alerts for high energy neutrino events now also initiate the HitSpool data read-out, as discussed in Section 5.2. The HitSpool data flow is shown as a dashed blue line. The EventBuilder handles standard data requests from IceCube's trigger system.

3.3 Overview of neutrino sources and background in IceCube
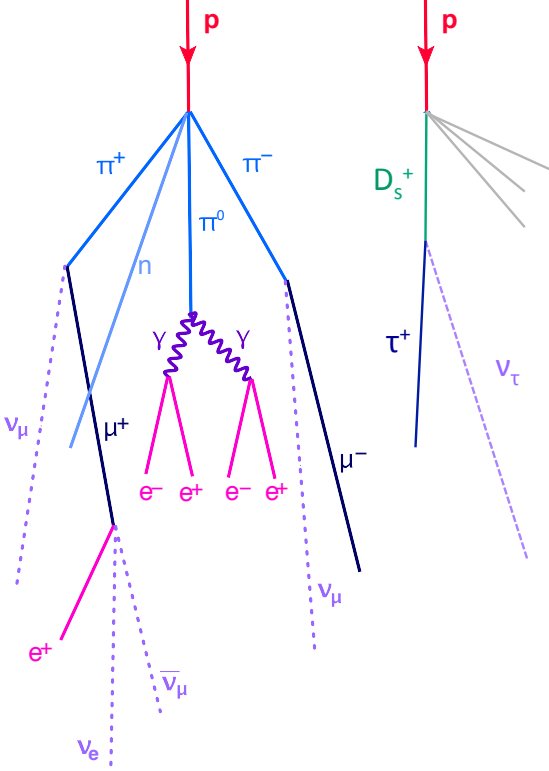
Figure 3.9: Sketch of air showers producing the conventional (left) and prompt (right) atmospheric neutrino flux.

The majority of event triggers in IceCube are caused by muons and neutrinos produced in the Earth's atmosphere. This so-called atmospheric flux is generated when cosmic rays interact with molecules in the upper atmosphere, thus creating aplenty pions, kaons and other particles in extensive air showers. These particles either undergo another interaction or they decay. The leptonic decay channel for charged pions and kaons, with a subsequent decay of the muon, reads:

$$\pi^{\pm}, K^{\pm} \rightarrow \quad \mu^{\pm} \quad + \quad \overset{(-)}{\nu}_{\mu}$$
$$\searrow$$
$$e^{\pm} + \overset{(-)}{\nu}_{e} + \overset{(-)}{\nu}_{\mu} \qquad .$$

The neutrino flux stemming from the above decay chain is named *conventional* and, except for the production through neutrino oscillation below 100 GeV neutrino energy, does not yield tau neutrinos. Tau flavor can, however, be produced in the decay of heavier particles, mainly D_s mesons. Prompt D mesons in general, with a lifetime on the order of picoseconds, generate a yet unobserved component called *prompt* atmospheric neutrino flux.

When muons in the above decay chain carry an energy greater than around 2.5 GeV, their decay length becomes larger than the distance they have to travel to reach the Earth's surface. Such atmospheric muons trigger the detector at a rate of 3000 Hz and are the predominant background to all neutrino searches.

Although cosmic rays generating the atmospheric fluxes were first detected almost a century ago, the mechanism accelerating their constituents to these high energies is still not settled. Every stable nuclei and particle could be found in cosmic rays, with protons and helium being most abundant contributing 79 % and 15 %, respectively [16]. From $\mathcal{O}(\text{GeV})$ to $\mathcal{O}(100\,\text{TeV})$, the cosmic ray composition is approximately constant. The all-particle energy spectrum shown in Figure 3.10 follows a power-law with $dN/dE \propto E^{-2.7}$ over most of the energy range. The *knee* of the spectrum occurs at around 4×10^{15} eV and refers to a softening to a power law index of -3. At around $10^{18.5}$ eV, the spectrum

26

steepens again producing another distinct feature called *ankle*. Below the ankle, cosmic rays are expected to originate from within our galaxy, while a dominating extragalactic component is assumed above.

Valuable input on the still open question on the production sites and acceleration mechanisms of high energy cosmic rays is provided by neutrinos originating from the same source. In contrast to the atmospheric component, this neutrino flux is called *astrophysical*. The energy spectrum as well as the flavor composition and spatial distribution of astrophysical neutrinos provides an excellent test for acceleration and neutrino production models. Chapter 4 summarizes the current knowledge of the astrophysical neutrino signal gathered from measurements in IceCube.
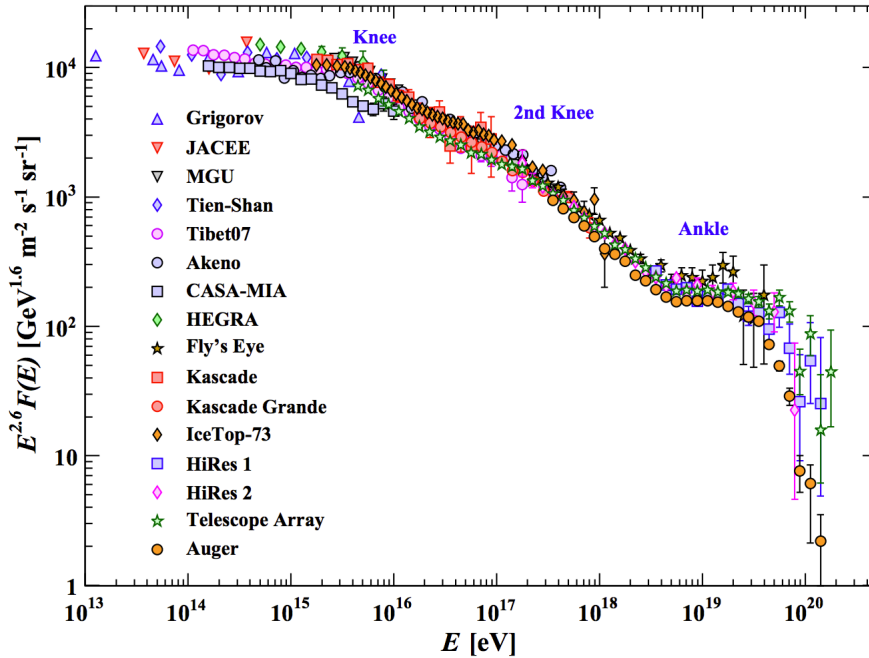
Figure 3.10: The cosmic ray all-particle energy spectrum [16], as measured by different air shower experiments, was multiplied by $E^{2.6}$ to make its features more prominent. Here, E ist the energy per nucleus of cosmic ray particles and $F(E)$ is the flux in units of $GeV^{-1}m^{-2}s^{-1}sr^{-1}$.

3.4 Event signatures at high energies
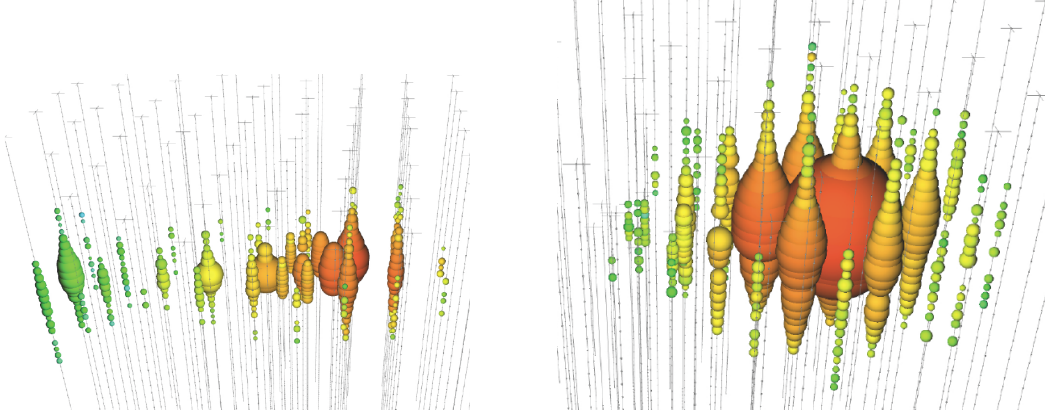
Figure 3.11: The two main event signatures seen in IceCube: muon tracks (left) and cascade events (right). Shown are event displays of simulated events in which the color coding represents the time flow from early (red) to late (blue) times and the size of a DOM illustrates the amount of light seen by it. For extremely high energy events, light from almost the complete IceCube extension of around 1 km is recorded in both cases.

In the energy regime this analysis targets, two main event signatures are well distinguishable: *tracks* and *cascades*. In addition to atmospheric muons, high energy muons are produced in ν_μ CC interactions (see Section 1.3). These muons leave a long Cherenkov light trail - a 200 GeV muon can travel roughly 1 km trough ice [39] - when traversing the detector, which is referred to as track.

This signature is different from the almost spherical pattern, called cascade, induced by any other neutrino interaction. Since the extent of a shower is only on the order of 10 m, as shown in Figure 3.12, IceCube is not able to resolve the shower details given the detector's coarser granularity. Despite intrinsic differences between electromagnetic and hadronic cascades and the ν_τ induced light pattern (see Chapter 2), a point-like light source is seen for all of them. The reconstruction of the interaction type in spite of the topologically highly similar light pattern of cascade events is the goal of this work.

It should be mentioned that more complex but not yet observed signatures, such as *double bang* events[3] , can be expected in IceCube for neutrinos at the PeV energy scale.

---

[3]Double bang dubs the two cascades created when a tau neutrino interacts via the CC channel, producing a first shower at the interaction vertex followed by a second cascade from the subsequently decaying tau lepton. The two showers are connected by the lepton track with an average length of around 50 m/PeV.
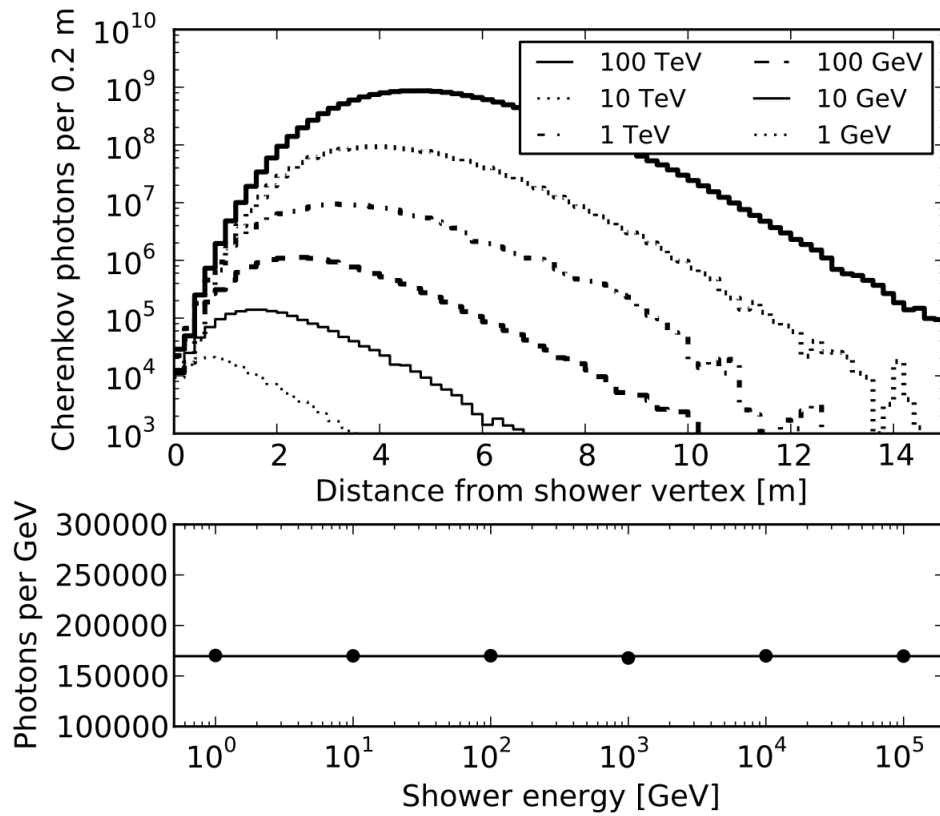
Figure 3.12: Longitudinal shower profile and energy dependence of the light deposition for electromagnetic showers in ice [34]. Since the cascade length is $\mathcal{O}(10\,\mathrm{m})$ and increases only logarithmically with the shower energy, IceCube's granularity does not allow for the resolution of the shower details.

4

The astrophysical neutrino signal in IceCube

In the investigation of astrophysical neutrinos, IceCube reported the first high energy excess over the atmospheric expectation already with its partial detector setup [40, 41]. This observation was followed by two PeV cascade events in the completed detector [42] and was finally confirmed by the detection of 28 high energy neutrino events rejecting a purely atmospheric origin at the $4\,\sigma$ level [43]. With the final detector configuration running since 2012, IceCube keeps reporting further high energy neutrino events [44–46], thus adding more and more puzzle pieces to our incomplete picture of the universe. Because the novel analysis discussed in this thesis builds on these results, the status is briefly summarized in this chapter.

4.1 The HESE event selection

IceCube's well-established HESE (*High Energy Starting Event*) event selection facilitated the first detection of extraterrestrial neutrinos [43]. The selection fights the atmospheric muon and neutrino background by requiring a minimum charge deposition of 1500 PE and by defining veto regions in the detector. To classify an event as a neutrino interaction with vertex within the fiducial volume, the event's first light must be recorded in this region and at most three of the first 250 PE are allowed to fall in the veto region (gray area in Figure 4.1). For a minimum required charge of 6000 PE, the selection rejects 99.999% of the atmospheric muon background (see Figure 4.2). The analysis of the astrophysical neutrino signal thus focuses on the dataset with the higher PE threshold implemented. In contrast to many IceCube analyses aiming at the extraction of ν_μ events due to their good angular resolution or focusing on events from the Nothern hemisphere to reject atmospheric muons, the HESE event selection achieves a similar sensitivity among neutrino flavors and directions [43].
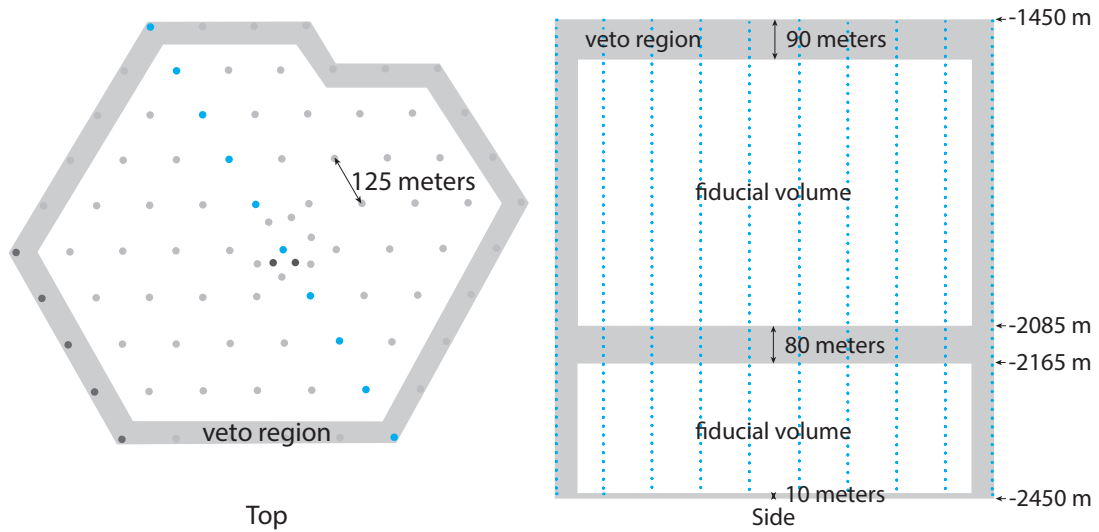
Figure 4.1: Pictured on the left is a top view of the IceCube detector with strings not yet deployed in the 2010-2011 season marked as dark gray dots. Strings colored blue in the left panel are shown in side view on the right-hand side. Gray shaded regions mark the veto region of the HESE event selection. Since a dominant fraction of the atmospheric muon background reaches the detector from above, a thick veto layer is required at the top. The central veto region removes the dust layer[1], which would allow horizontal muons to enter the fiducial volume without being identified by the outer vertical veto bands [43].

---

[1]The dust layer is a region at a depth between 2000 m and 2100 m with a high dust concentration and thus high light absorption.
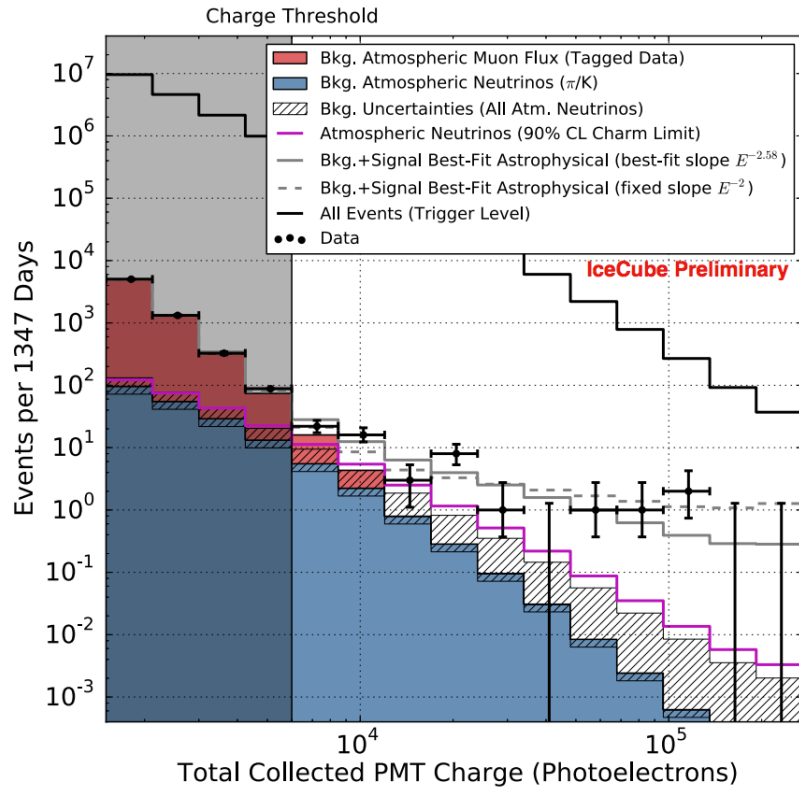
Figure 4.2: Charge distribution of events meeting the HESE criterion. Data shown in this plot were taken from 2010 to 2014 and illustrate the excess of high energy events above the atmospheric background prediction [45].

4.2 Flavor composition

Nuclei accelerated in gamma-ray bursts, supernova remnant shocks, starburst galaxies or other violent astrophysical events interact with close-by matter or radiation producing pions which, in analogy to the processes in an atmospheric air shower, create neutrinos and muons. Since these muons will most likely decay, a ratio of $1 : 2 : 0$ in $\nu_e : \nu_\mu : \nu_\tau$ is expected at the production site which mixes to an approximately equal flavor distribution of $1 : 1 : 1$ at Earth via long baseline neutrino oscillation (see Section 1.1).

However, different neutrino production scenarios can be imagined [47]: For muon-damped cases, in which the muons are absorbed before they decay, a flavor composition of $0 : 1 : 0$ is expected at the source. If astrophysical neutrinos are instead produced in the decay of neutrons generated via photo-disintegration of nuclei, an initial flavor ratio of $1 : 0 : 0$ is predicted. Yet another model, valid at high energies, sees neutrinos produced in semileptonic decays of charm quarks, yielding a $1 : 1 : 0$ ratio at the source. Table 4.1 summarizes the model-dependent flavor composition at the astrophysical source as well as the expected ratio after long baseline oscillation.

A global fit to IceCube's high energy events, selected within various analyses, demonstrates that ν_μ CC events can be well differentiated from cascades by their track-like topology, but that the event signature of ν_τ CC events is so far indistinguishable. Searches for ν_τ CC interactions have not succeeded yet [23, 24], leaving the lower ν_τ fraction in IceCube's flavor triangle, shown in Figure 4.3, unconstrained. This emphasizes the need for a method able to identify shower types, even if a discrimination by a distinct event topology, such as the ν_τ's double bang signature, is not possible. This work can provide valuable input in constraining the ν_τ and ν_e contribution in shower-like events.

Due to the large number of track-like events in the combined data sample of the global fit, neutrino production solely through neutron decays can be excluded. Other neutrino production scenarios are still in agreement with the best-fit results.

ν production	$(\nu_e : \nu_\mu : \nu_\tau)^{\text{production site}}$	$(\nu_e : \nu_\mu : \nu_\tau)^{\text{Earth}}_{\text{NH}}$	$(\nu_e : \nu_\mu : \nu_\tau)^{\text{Earth}}_{\text{IH}}$
pion, muon decay	$1 : 2 : 0$	$1.02 : 1.01 : 0.97$	$0.92 : 1.06 : 1.02$
muon-damped	$0 : 1 : 0$	$0.23 : 0.39 : 0.38$	$0.18 : 0.44 : 0.38$
neutron decay	$1 : 0 : 0$	$0.56 : 0.23 : 0.21$	$0.56 : 0.18 : 0.26$
semileptonic charm decay	$1 : 1 : 0$	$0.79 : 0.62 : 0.59$	$0.74 : 0.62 : 0.64$

Table 4.1: The production site and oscillated flavor composition of the astrophysical neutrino flux for different production scenarios and neutrino mass hierarchies. The values were calculated using Equation 1.3. Variations due to the uncertainty in the neutrino mixing parameters are on the order of $10\,\%$ [8].
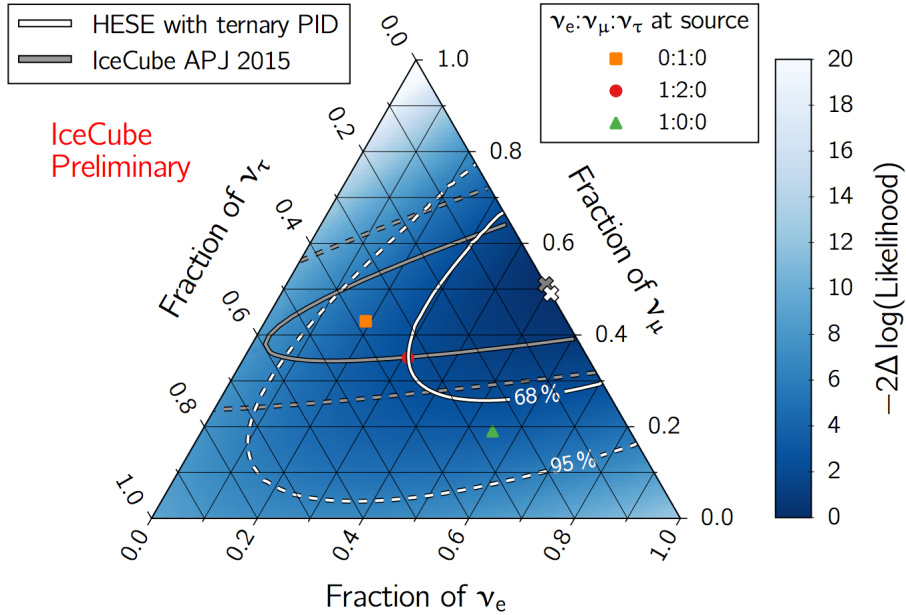
Figure 4.3: The best-fit flavor composition and confidence regions of high energy events in IceCube [24], obtained from a global fit only distinguishing track and shower topologies (gray marker and lines) [48], and an analysis additionally exploiting characteristic features of ν_τ events (*ternary PID*, white marker and lines) [24]. The orange, red and green marker shows the flavor composition expected at Earth for different flavor ratios at the source. Axes must be read along the tick lines.

4.3   Energy spectrum

First order diffusive shock acceleration[2] is a favored model to explain how cosmic rays are accelerated to the high energies observed at Earth. In the test-particle approximation, this acceleration mechanism yields an energy spectrum at the source following $E^{-\gamma}$, with $\gamma = 2$. Taking into account the energy dependent time, after which cosmic rays escape our galaxy, $\tau_{esc} \propto E^{-\delta}$, with $\delta \approx 0.6$, the galactic cosmic ray energy spectrum is predicted to follow an inverse power law with a spectral index of 2.6 [51]. This in good agreement with the measured spectrum proportional to $E^{-2.7}$.

In the standard source scenario, pions and muons, generated by the accelerated nuclei, decay without losing much energy. Astrophysical neutrinos are thus expected to

---

[2]E. Fermi developed the concept of charged particle acceleration through multiple collisions with moving magnetic fields in interstellar gas clouds [49]. In this model, the energy gain is proportional to the squared particle velocity, and the mechanism is thus called *second order Fermi acceleration*. The extension of the concept to particle acceleration in astrophysical shock waves [50] yields a linear energy gain-particle speed correlation and is named *first order Fermi acceleration* or *first order diffusive shock acceleration*.

inherit the cosmic ray energy spectrum with $\gamma \approx 2$ at the production site. The astrophysical neutrino energy spectrum has been measured by various analyses in IceCube, delivering still consistent, but sample dependent results. The best-fit spectral index to IceCube's first 28 high energy neutrinos is 2.2 ± 0.4 [43]. Due to the absence of detected PeV events in the following years, a more and more softer energy spectrum, with a spectral index of 2.3 ± 0.3, 2.58 ± 0.25 and $2.92^{+0.33}_{-0.29}$ for the three year, four year and six year continuation of the analysis describes the HESE data best [44–46]. A consistent result of $\gamma = 2.48 \pm 0.08$ is obtained from a cascade sample containing showers from $\mathcal{O}(1\,\mathrm{TeV})$ to $\mathcal{O}(200\,\mathrm{TeV})$ [52]. However, these results, which stem from cascade-dominated event samples, are in an approximately $2\,\sigma$ tension with the best-fit spectral index of 2.19 ± 0.10 from an IceCube analysis studying track-like ν_μ events from the Northern hemisphere [53] (see Figure 4.4).

This, of course, raises the question whether the astrophysical neutrino flux exhibits spectral and directional features. With the current statistics at hand, it is, however, not yet possible to e.g. significantly favor a two-component energy spectrum over the single power law [46, 52] or provide evidence for the existence of a spectral cut-off [48, 53]. A global fit to the cosmic neutrino flux in IceCube, performed in 2015, already disfavored the unbroken power law hypothesis with $\gamma = 2$ at the $3.8\,\sigma$ level [48]. An imminent updated global fit [54] and accumulating detected high energy events may align our understanding of the astrophysical neutrino spectrum.
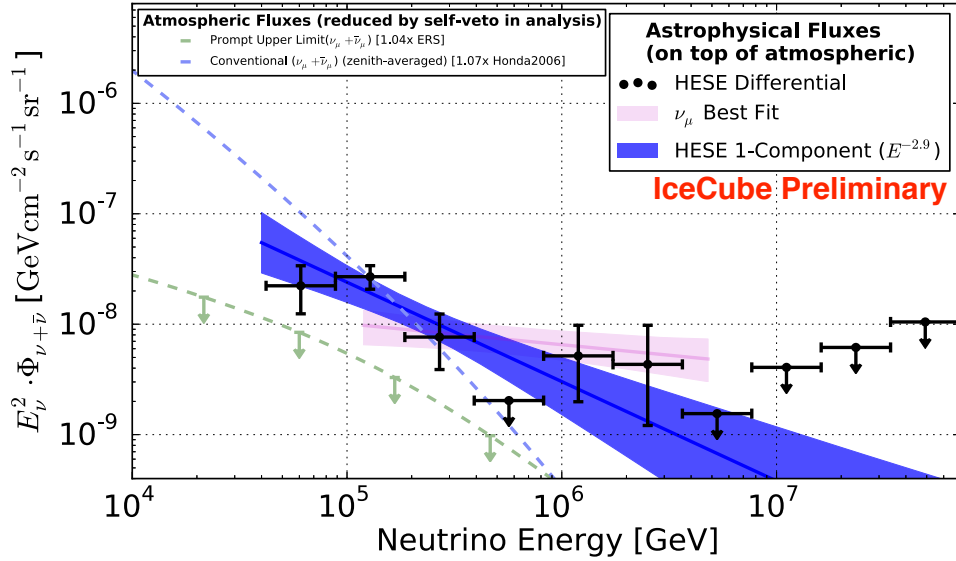
Figure 4.4: The best-fit energy spectrum to six years of data collected in IceCube's HESE event selection (black points, atmospheric components are subtracted) is shown as a blue solid line [46]. The blue band marks the fit's $1\,\sigma$ uncertainty range. The pink band shows the best-fit spectrum from the ν_μ analysis together with its $1\,\sigma$ uncertainty, and also marks the analysis' sensitive energy range [53].

4.4 Spatial distribution

Retracing the cosmic neutrino flux to a specific astrophysical point-source would be an experimental breakthrough. With each additional year of data, events selected by the HESE criterion are therefore tested for spatial clustering. The test statistic (TS), quantifying the spatial clustering, is defined as the logarithm of the ratio of the maximum likelihood under the point-source hypothesis over the maximum likelihood under the isotropic flux hypothesis[3]. Comparing the TS distribution observed in the data to skymaps with event directions scrambled in right ascension allows for a statement on the significance of the clustering in the data. Figure 4.5 shows the most recent skymap published by the IceCube collaboration. With the diffuse flux assumption as null hypothesis, the p-value[3] for the six year event sample is 77% [46]. Excluding track-like events, which could dominate the likelihood due to their superior angular resolution of $\lesssim 1°$, yields a p-value of 44%. There is thus no evidence for neutrino point-sources in the high energy data collected so far.
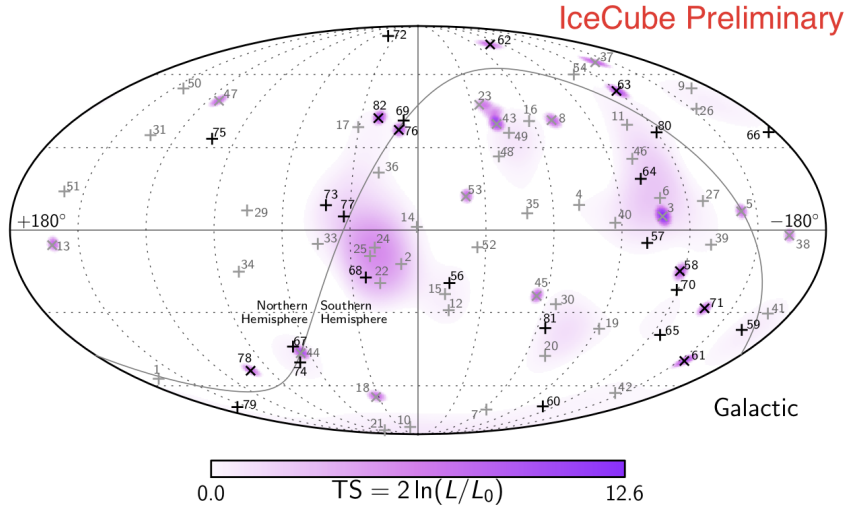
Figure 4.5: Event directions of showers ($+$) and tracks ($\times$), selected by IceCube's HESE stream, in galactic coordinates [46]. The figure includes the most recent results up to May 2016, with events recorded between May 2014 and May 2016 marked black. The color scale refers to the test statistic (TS) of the point-source clustering test.

---

[3] See Section 7.3 for details on hypothesis testing via the likelihood-ratio approach.

5

Cascade type identification in IceCube

Chapter 2 explained how the particle content of a hadronic shower differs from the electromagnetic cascade constituents. Due to its coarse granularity designed for neutrino detection at the highest energies, IceCube is, however, not equipped to geometrically resolve the shower composition. Nevertheless, certain particles can be indirectly identified via the delayed light signal, or *echo*, that they induce in the detector. In this chapter, the echo approach, sources for delayed photons and neutrons as the main channel investigated in this work will be presented. The chapter concludes by describing the experimental implementation of the echo recording for high energy events.

5.1 Exploiting the cascade echo

This analysis focuses on the indirect detection of neutrons as an approach to a possible shower type distinction in IceCube. When neutrons are captured by nuclei in the ice, delayed photons are produced when the excited nuclei return to their ground state. Since neutrons are more abundant in hadronic than in electromagnetic cascades, the neutron capture occurrence is consequently much higher for hadronic showers and the corresponding echo strength allows for an identification of the shower type.

5.1.1 Neutrons in hadronic cascades and the neutron echo

As Figure 5.1 shows[1], around 60 % of all the neutrons produced[2] in a particle shower in ice get captured. With around 98 % of the neutrons capturing on hydrogen [55], this element is the predominant neutron absorber. The hydrogen nucleus together with the captured neutron form an excited deuterium state, which gives off a 2.223 MeV gamma quantum upon de-excitation. This photon subsequently generates on average 14 Compton electrons, which produce a detectable Cherenkov signal as long as the energy of the created electrons and positrons exceeds the Cherenkov threshold of 0.26 MeV.

In hadronic showers, neutrons are not only produced by high energy particle interactions at the start of the cascading, but also via evaporation[3] at later stages, when particle

---

[1]Plots in this section show simulation results produced with the Geant4 simulation toolkit (see Chapter 6 for details).

[2]All neutrons produced in inelastic neutron scatterings are counted as *new* neutrons in Geant4.

[3]The use of the term *evaporation* here goes back to Victor Weisskopf, who mathematically described the emission of neutrons from excited nuclei similar to the models for the evaporation of particles from low temperature bodies.
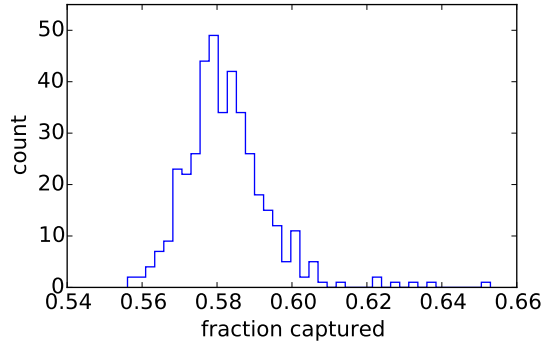
Figure 5.1: The fraction of neutrons captured, i.e. the ratio of the number of neutron captures over the total number of neutrons in a cascade, is shown for 400 showers in ice induced by $1\,\text{TeV}$ π^+ particles.

energies are on the order of the nuclear binding energy.

After creations, neutrons scatter down to thermal energies. With $\mathcal{O}(100\,\text{barn})$ [56], the neutron scattering cross section is large and thermal energies are reached within about $1\,\mu\text{s}$ [57]. The thermalized neutrons subsequently perform a random walk until they are captured. The microscopic absorption cross section for a component k in a material is given by [58]:

$$\sigma_{ab,k} = \frac{\pi}{2}\,\sigma_{ab,k}^{2200}\,\sqrt{\frac{293\,\text{K}}{T_N}} \qquad , \tag{5.1}$$

with the neutron temperature T_N and σ_{ab}^{2200} being the absorption cross section for neutrons at a velocity of $2200\,\text{m/s}$. Given the microscopic cross section σ_{ab}, the macroscopic absorption cross section for a material with atomic number density N reads:

$$\Sigma_{ab} = \sigma_{ab} \cdot N = \sigma_{ab} \cdot \frac{\rho \cdot N_A}{M} \qquad , \tag{5.2}$$

where ρ is the material density and M the molecular weight. With the neutron velocity v, the probability for a neutron capture P_{cap} in time interval dt is given by:

$$P_{cap}(v) \propto \sum_k \Sigma_{ab,k}(v) \cdot w_k \cdot v\ dt = 1/\tau_{cap}\ dt \qquad . \tag{5.3}$$

Therein, w_k is the stoichiometric fraction of component k.

As the microscopic absorption cross section is inversely proportional to the neutron velocity over the relevant energy range for H_2O, the capture time $\tau_{cap} \propto \sigma_{ab}(v) \cdot v$ is a constant. The time distribution of neutron capture occurrences in a given material is

thus an exponential function with decay time

$$\tau_{cap} = \frac{1}{\sum_k v \cdot \sigma_{ab,k}(v) \cdot w_k \cdot \frac{\rho \cdot N_A}{M}} \qquad . \tag{5.4}$$

Ice at IceCube's depth has a density of around $920 \, \text{kg/m}^3$ [59] and a molecular weight of $M \approx M_{H_2O} \approx 18 \times 10^{-3} \, \text{kg/mol}$. The neutron absorption cross section of hydrogen is $0.3326 \, \text{barn}$, thus yielding the constant product $\sigma_{ab} \cdot v = 0.3326 \, \text{barn} \cdot 2200 \, \text{m/s} = 7.317 \times 10^{-26} \, \text{m}^3/\text{s}$. When neglecting the contribution from oxygen, which has an approximately 1000 times smaller absorption cross section than hydrogen, the expected neutron capture time in IceCube is $222.2 \, \mu\text{s}$. The Cherenkov photon time spectrum in Figure 5.2 illustrates the exponential behavior.

As shown in Figure 5.3 and Figure 5.4, neutrons are captured on average $4 \, \text{m}$ distant from the interaction vertex in a $1 \, \text{TeV}$ shower and travel around $10 \, \text{cm}$ from their point of creation before being captured. The $2.223 \, \text{MeV}$ photon, which is generated in the de-excitation process, travels roughly $50 \, \text{cm}$ in ice and induces on average approximately 230 Cherenkov photons.
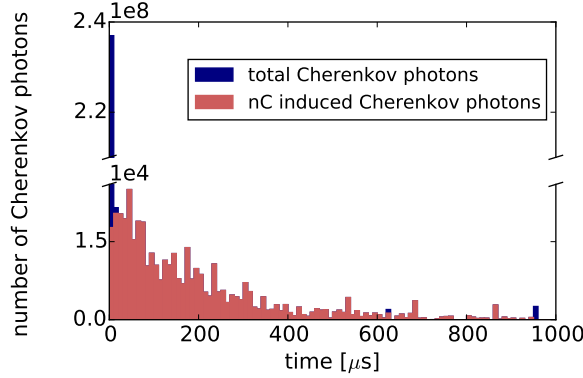
Figure 5.2: The time distribution and the neutron capture (nC) induced component of the Cherenkov signal is exemplarily shown for a cascade produced by a $1 \, \text{TeV} \, \pi^+$ interacting in ice. The delayed signal originates predominantly from neutron capture processes. The large fluctuations in the delayed time spectrum are caused by the discrete neutron capture times.
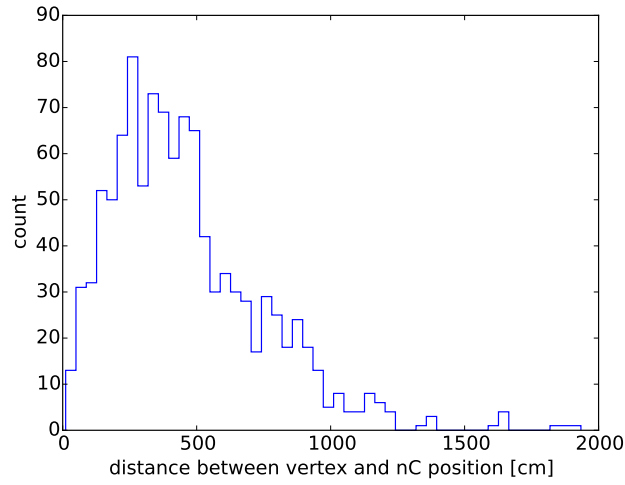
Figure 5.3: The distance between the event vertex and the neutron capture positions is exemplarily shown for a 1 TeV shower induced by a π^+.
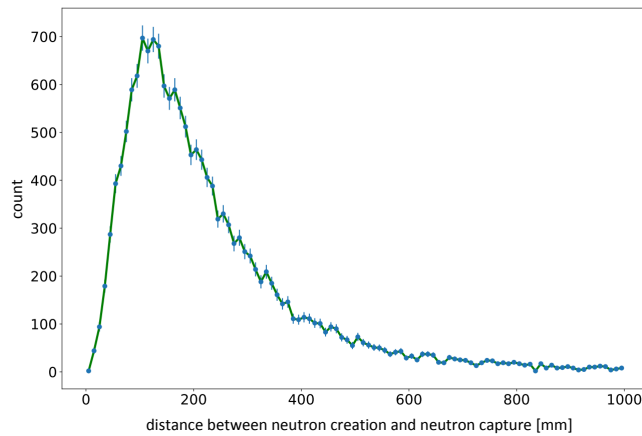
Figure 5.4: The distance between the point of the neutron creation and the neutron capture position [60]. The peak of the distribution is located at around 115 mm.

Compared to the total number of Cherenkov photons in a hadronic shower, the delayed neutron capture signal makes up around 0.1 % (see Figure 5.5). As the electromagnetic component of a hadronic cascade increases with shower energy, the ratio is also energy dependent. Figure 5.5 furthermore demonstrates that electromagnetic cascades contain a small hadronic component e.g. from photonuclear reactions, as the fraction of neutron capture induced photons is around 7×10^{-3} %.

40
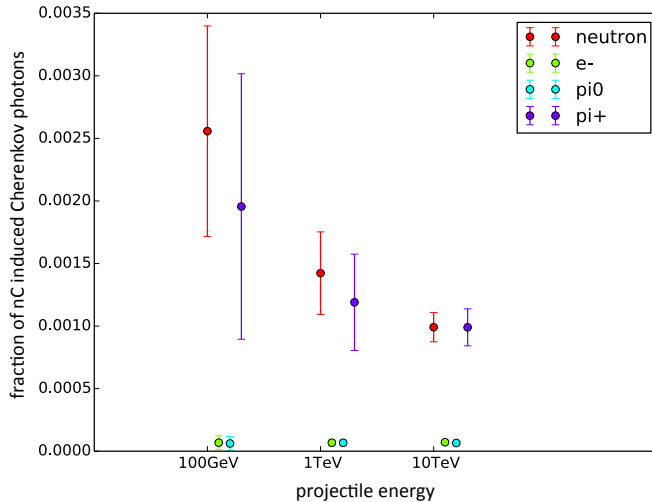
Figure 5.5: The fraction of neutron capture induced Cherenkov photons over the total Cherenkov photon count for different projectiles. The error bars represent the spread of the hadronic fluctuations; the uncertainty on the mean is between 5 % and 10 %.

5.1.2 Further sources of delayed photons

A delayed light signal can also arise from sources other than neutron capture processes. In the following section, additional contributions are discussed particularly in regard to a shower type analysis.

The muon echo

Charged pions are copiously produced in hadronic reactions. After losing their kinetic energy via ionization, π^- capture on nuclei while π^+ decay to positively charged muons. With a lifetime of approximately $2.2\,\mu s$, the thus produced muons will decay into positrons with an energy of around $35\,\text{MeV}$. These positrons are, in analogy to the photon following a neutron capture, the origin of a detectable delayed Cherenkov signal called *muon echo*. Although the intensity of the muon echo is roughly five times larger than that of the neutron echo [25], its time profile makes an utilization in a shower type analysis difficult. With a characteristic time of around $2\,\mu s$, the muon echo falls in the midst of the PMT afterpulses (see Section 3.2.1), whose intensity of around 6 % of the prompt Cherenkov signal is comparatively high. Since the afterpulsing behavior differs from DOM to DOM and is not precisely known for the individual modules, the extraction of the muon echo seems very difficult to implement and, if at all, will only be possible in a statistical sense. Figure 5.6 demonstrates how the echo is covered by the afterpulses in a $1\,\text{PeV}$ hadronic cascade. In addition, prompt Cherenkov photons that travel long distances can have delay times of $\mathcal{O}(\mu s)$ and thus also overshadow the muon signal.

41
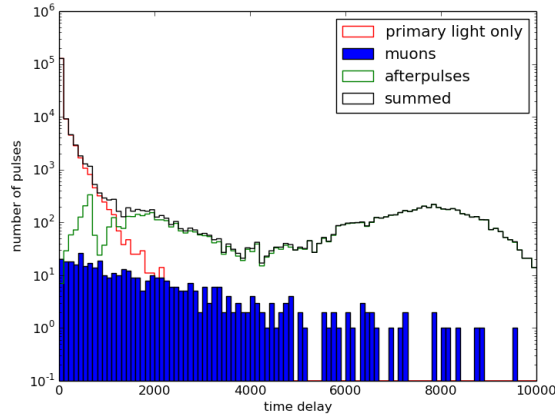
Figure 5.6: The primary light, muon echo and afterpulses in the DOM closest to the vertex of an approximately 1 PeV hadronic cascade as simulated in a Toy Monte Carlo study [61]. Detection artifacts other then the PMT afterpulsing are not included. The time delay is given in units of nanoseconds.

Radioluminescence

Luminescence encompasses all light emission processes of a material that are not caused by heat. Ionizing particles passing a substance, for instance, can excite atomic electrons in the material. When the atom returns to its ground state, it loses the extra energy via the emission of a photon. This process is called radioluminescence and its intensity and decay time strongly depend on the material itself and its thermal state. Luminescence of ice is not very well explored in general and measurements at IceCube's pressure and temperature conditions are missing in particular. Ongoing efforts within the IceCube collaboration to measure the phenomenon found that the intensity of radioluminescence can be up to 10 % of the prompt Cherenkov signal for PeV energy events [62]. The so far measured intensities for α particle induced luminescence in water and ice range from $0.2\,\gamma/\mathrm{MeV}$ to $20\,\gamma/\mathrm{MeV}$ [62] and decay times vary from tenths and hundreds of nanoseconds for crystalline ice [63] to seconds for amorphous ice [64]. These numbers can again be very different when considering impurities as well as temperature and pressure conditions in the Antarctic ice. An in-situ measurement using the 1750 m deep *SPICEcore* hole[4] close to IceCube is under preparation [65].

Spallation products

Spallation refers to the breakup of nuclei when bombarded with high energy particles. The daughter nuclei produced in spallation processes can be unstable and return to their ground state predominantly through a beta decay. Given the corresponding iso-

---

[4]The *SPICEcore* (South Pole ICE core) is an ice sample drilled out of the Antarctic ice. Ice cores are commonly used to study the history of climate and environment conditions.

tope lifetimes, the delayed emission of beta radiation leads to a delayed Cherenkov signal. Since spallation strongly contributes to the background of low energy analyses in neutrino observatories, it has been amply studied in [66] and [67] for the Super-Kamiokande experiment, which uses water as detector material. These investigations found that, although generating little or no Cherenkov radiation themselves, charged pions, neutrons and photons are the most important isotope producers via the breakup of oxygen. However, delayed light emission due to nuclear spallation plays a negligible role in this analysis, as Figure 5.2 shows that the vast majority of delayed photons originates from neutron captures. Relevant isotope yields and lifetimes are too low and long[5], respectively, to impact the delayed photon spectrum.

Correlated noise

Figure 3.7 shows that the chronology of correlated noise can by chance mimic the exponential shape of the neutron capture signal. Per DOM and millisecond, between 0.28 and 0.34 PE are expected from this noise component (see Section 3.2.1). However, in contrast to the neutron echo, the start of a sequence of correlated noise hits is neither associated with the temporal and positional detection of a neutrino event nor correlated between the optical modules.

Coincident atmospheric muons

Atmospheric muons depositing light within the time span and close to the DOMs the neutron echo is searched for can be falsely identified as belonging to the delayed signal. However, the additional rate per DOM due to atmospheric muons is small and ranges from around $3\,\mathrm{Hz}$ to $30\,\mathrm{Hz}$, depending on the module's location [37]. This corresponds to 0.003 to 0.03 hits from atmospheric muons in a millisecond time window, which is a negligible effect.

Afterpulses

The afterpulsing behavior of IceCube PMTs (see Section 3.2.1) has been measured up to $11\,\mu$s after incident illumination. The possibility of much later, so far unmeasured, afterpulses in coincidence with the neutron echo will be briefly discussed here. The delay time Δt of an afterpulse with respect to the prompt light signal is proportional to the mass-to-charge-ratio, m/q, of the ion it can be attributed to:

$$\Delta t \propto \sqrt{\frac{m}{q\,V_0}} \quad ,$$

where V_0 is the voltage between photocathode and first dynode. According to calculation by Hamamatsu, methane produces the peak at around $2\,\mu$s, while peaks in the 6-8 μs

---

[5]The largest atmospheric muon induced isotope yields were found for ^{16}N ($18 \times 10^{-7}\,\mu^{-1}\mathrm{g}^{-1}\mathrm{cm}^2$), ^{8}Li ($13 \times 10^{-7}\,\mu^{-1}\mathrm{g}^{-1}\mathrm{cm}^2$) and ^{12}B ($12 \times 10^{-7}\,\mu^{-1}\mathrm{g}^{-1}\mathrm{cm}^2$). The corresponding isotope half-lives are $t_{1/2,^{16}\mathrm{N}} = 7.13\,\mathrm{s}$, $t_{1/2,^8\mathrm{Li}} = 0.838\,\mathrm{s}$ and $t_{1/2,^{12}\mathrm{B}} = 0.0202\,\mathrm{s}$ [66].

range of the afterpulsing spectrum stem from caesium, although a definite assignment of peaks and production candidate ions is not possible [68]. However, this allows one to conclude that ions with an around 100 times higher mass-to-charge-ratio than methane would need to be present to generate afterpulses after $30\,\mu s$, which marks the start of the time window in which the neutron echo is searched for. The possibility of afterpulses mimicking or covering the neutron echo can thus be excluded.

5.2 HitSpooling for HESE events

Photons originating from neutron captures arrive at the sensors with a delay of up to $\mathcal{O}(1\,\mathrm{ms})$ with respect to the neutrino interaction time, which is outside the standard readout out windows of around $10\,\mu s$ in IceCube. Therefore, HitSpool data (see Section 3.2.2) for high energy cascade events are needed.

In order to extract high energy cascades from the data sample, this work applies IceCube's HESE event selection (see Section 5.2). When an interesting event is identified by the HESE filter, a request for HitSpool data is sent to the HitSpool request daemon (see Section 3.2.2). This initiates the collection of HitSpool data in a time window of $\pm 0.5\,\mathrm{s}$ around the trigger time, which will be sent to the data warehouse in the North via satellite. A HESE alert with a charge deposition greater than $6000\,\mathrm{PE}$ is expected approximately once per month and results in a raw data size of around $50\,\mathrm{MB}$ in compressed format to be transferred. Upon arrival in the North, the compressed raw data is unpacked and converted into IceCube's commonly used *I3* file format[6], in which the $1\,\mathrm{s}$ data chunk will be subdivided into *frames*, each holding the information on a $100\,\mu s$ block of hits[7]. These files are the input of the processing chain in IceCube's software framework *IceTray*. Therein, further trigger algorithms are run and hits are marked with a trigger information tag. Data are subsequently passed to a processing chain for noise cleaning, event identification and event reconstruction:

HiveSplitter: *HiveSplitter* [69, 70] aims at the assignment of hits belonging to a physics event or being noise and separating hits from coincident events. The algorithm works on the basis of the individual active volume (IAV), which is a time dependent spatial volume calculated for each hit DOM, respecting the detector geometry (see Figure 5.7). The size of a hit's IAV is the volume in which further hits can be causally connected to it, i.e. originate from the same (sub)event and is calculated based on causality arguments respecting the spatial separation of the DOMs and particle and photon propagation. Four overlapping IAVs define the minimal set of clean hits. Further hits will be added to that cluster if they fulfill the overlapping criterion.

---

[6]I3 files are portable binary archives.

[7]Usually, one frame contains the hit information of one event. This concept was adjusted to fit the storage requirements of HitSpool data continuously recorded over a longer time span.
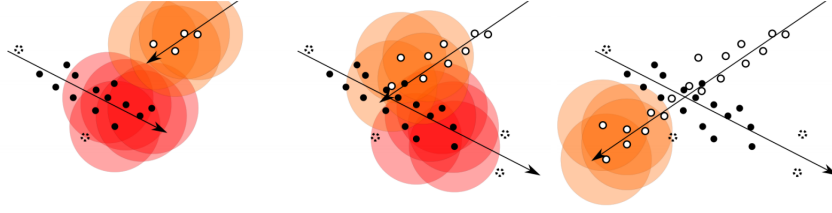
Figure 5.7: The HiveSplitter algorithm is able to separate the hits (black and white filled circles) of coincident events, as shown here for two muon tracks (arrows), by looking for overlaps in the IAVs of hits (red and orange spheres). Hits whose IAV do not connect with any cluster are assumed to be noise (dashed circles) [69].

NoiseEngine: The NoiseEngine module [71] is run on hit clusters identified by HiveSplitter and checks whether they were produced by noise or physics events. To decide on this, the directionality of vectors connecting all possible pairs of hit DOMs is investigated. For physics events, the direction of these vectors will be in line with the direction of the event itself, while no directionality is expected for noise clusters.

Subthreshold Muons: Low energetic atmospheric muons, which cannot be identified by any trigger, contaminate the raw data recorded by the HitSpool system. A routine [38] seeking to identify these so-called *subthreshold muons* is thus implemented in the HitSpool processing chain. Hits will be associated with an atmospheric muon event when they either cause a corresponding trigger or when they form a (non-triggered) cluster according to the HiveSpiltter algorithm and are subsequently classified as not being noise by the NoiseEngine module.

Monopod: As a likelihood-based event reconstruction algorithm, *Monopod* looks for the event hypothesis compatible with a measured hit pattern by maximizing a corresponding likelihood function [34]. The algorithm obtains the light expectation of a given source configuration from spline-interpolated photon simulation results [72, 73], so-called *photontables*. As Monopod includes a noise contribution in its light yield hypothesis, the algorithm is run on uncleaned pulsemaps containing HLC and SLC pulses.

In addition to the I3 format, data is available in hit tables in the *hierarchical data format* (HDF5), which hold information on the location, time, LC condition and trigger and noise contribution for each hit.

The link between real-time HESE alerts and the HitSpool system was implemented in February 2016; since February 25, 2016, the automated HitSpool data taking for HESE events has been running in a stable fashion. Table 5.1 summarizes the statistics for the HESE HitSpool data.

event charge	expected event rate	dataload per event	events recorded
1500 - 6000 PE	4 per day	$\approx 50\,\mathrm{MB}$	2674
>6000 PE	1 per month	$\approx 50\,\mathrm{MB}$	34

Table 5.1: HESE HitSpool data recorded from 02/25/2016 - 12/31/2017.

6

Event simulation

To study how the delayed light signal will reveal itself in IceCube data and to have a reference sample to which experimental data may be compared, simulated events are needed. The event simulation must encompass the full modeling from the neutrino interacting in the Antarctic ice to the final signal reported by the detector. The tools designated for the simulation of the different event stages will be presented in this chapter.

6.1 Simulation of neutrino interactions with PYTHIA 8 and DIRE

The neutrino DIS in ice is simulated as a collision of high energy neutrinos with protons at rest by means of the *PYTHIA 8* program [74] - a toolkit designated to model final state particles from high energy collisions. Per se, PYTHIA 8 does not provide the simulation of lepton-hadron interactions [74], although DIS is implemented as a process with the limitation of missing associated parton showers [75]. In this work, this is compensated by the *DIRE (DIpole REsummation)* parton shower algorithm [76], which is available as a plugin for PYTHIA 8. The versions used here are PYTHIA 8.2.15 and DIRE 0.900.

The complete event and final state particle information is written to *HepMC* files [77]. From these event records, the particle information is collected and translated to the input format required by the subsequent simulation packages.

6.2 Particle tracking with Geant4

Geant4 (Geometry and Tracking 4) [78–80] is a commonly used toolkit for the simulation of particle propagation through matter. It provides the implementation of complete detector setups and particle tracking for energies from $\mathcal{O}(100\,\mathrm{eV})$ to the TeV scale. In this analysis, GEANT4 release 10.01 patch 01 is used to simulate the passage of particles produced in the neutrino DIS through ice.

Geant4 based simulations were run in two ways:

1. Stand-alone Monte Carlo that takes into account the ice properties and calculates the Cherenkov photon yield, but does not model the IceCube detector and its response.

2. Full IceCube Monte Carlo that tracks Cherenkov photons using IceCube routines and fully describes the detector response.

6.2.1   Target material

At this stage of the simulation, no detector details are necessary and the detector setup is a large volume of ice in the Geant4 stand-alone as well as in the full detector simulation[1]. The isotope composition of the implemented ice is deduced from Antarctic snow precipitation [59]. An air content of $0.09\,\mathrm{cm}^3/\mathrm{g}$ [81] at IceCube depths is assumed. In terms of mass fractions, the natural medium is composed of:

- 99.9892 % Ice:

  - 99.7 % H_2O
  - 0.03 % HDO
  - 0.2 % $H_2{}^{18}O$

- 0.0108 % Air
  (implemented through *G4Air* from the Geant4 Material Database)

In the event simulation, the air content was overestimated by assuming a mass fraction of 8.3 %[2]. While this overestimation had no effect on the Cherenkov photon counts, the capture time increased from around $190\,\mu\mathrm{s}$ to $217\,\mu\mathrm{s}$ with the corrected detector material (see Figure 6.1), which is compatible with the expected capture time of around $220\,\mu\mathrm{s}$ derived in Section 5.1.1.

6.2.2   Geant4 settings

Geant4 allows one to tune simulation settings to suite an application's specific requirements. Settings found important in this work are discussed in the following paragraphs.

Physics lists

The most fundamental setting the user can control is the choice of physics processes and models used in the simulation. The implementation of the user selection happens via so-called *physics lists*. Geant4 provides predefined lists, called *Reference Physics*

---

[1]In the Geant4 stand-alone simulation, a 500 m radius-sphere is implemented while a $10\,\mathrm{km}\times10\,\mathrm{km}\times6\,\mathrm{km}$ box is modeled in the full detector simulation (see Section 6.3). Both setups are chosen large enough to fully contain the simulated cascades.

[2]which is the corresponding volume fraction.
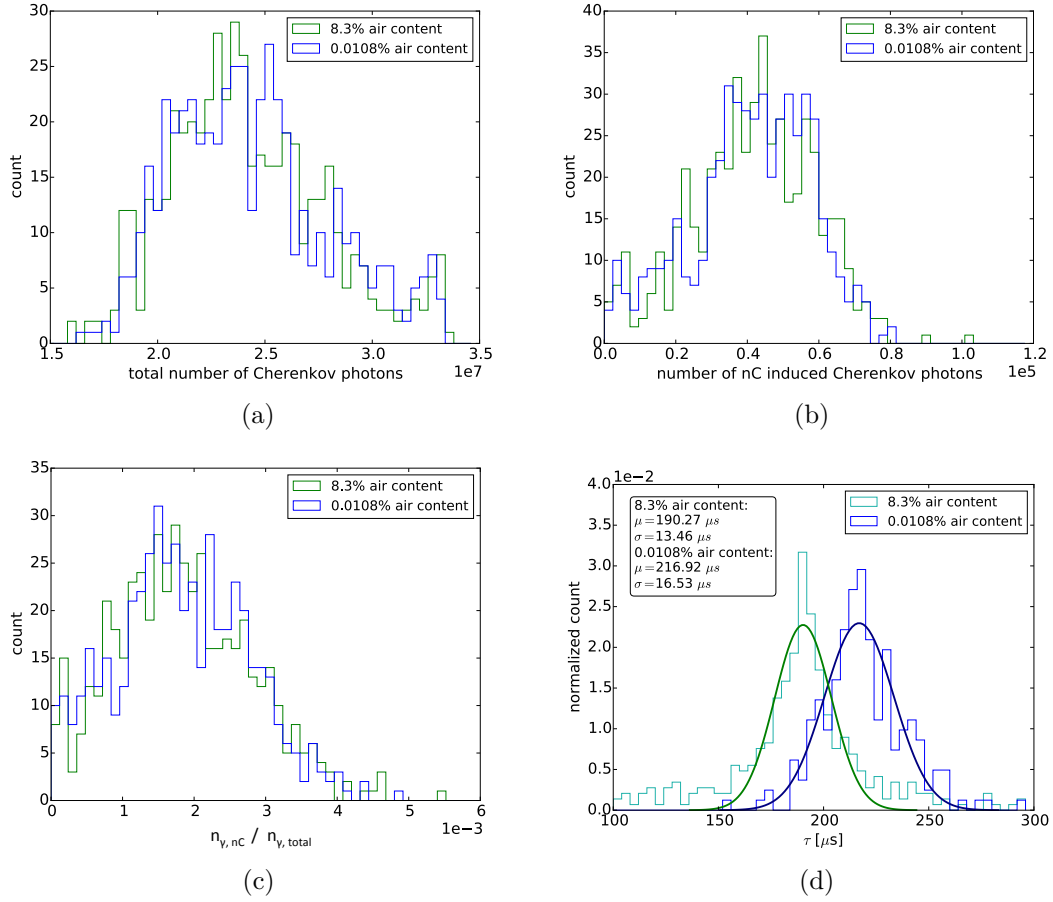
Figure 6.1: The impact of the initially overestimated air content. 500 hadronic showers induced by π^+ primaries enter each histogram in subfigures (a)-(c). The histograms in subfigure (d) contain the fitted decay times of multiple showers initiated with different primary types and energies.

Lists, which conveniently register all the processes and models most suitable for a certain application. This analysis uses the reference physics list *QGSP_BERT_HP* (see Appendix A), recommended for the observation of particle fluxes from high energy collisions in general and furthermore incorporating the high precision neutron model for the adequate modeling of the corresponding neutron fluxes.

However, in the early stages of the analysis, a modified version of the Geant4 build-in example *Hadr06* physics list (from now on called *modHadr06*) was used (see Appendix A). In terms of neutron production, the relevant difference is the implementation of the *FTFP* model for high energy hadronic collisions in *modHadr06*, while *QGSP_BERT_HP* uses the *QGSP* model. A comparison of results from *modHadr06*

to *QGSP_BERT_HP*, also with regard to further Geant4 settings, will be given at the end of this section.

Accuracy settings

Geant4 furthermore allows the user to tune parameters that control the precision of the simulation. The *range cut*, *step size* and the *database binning* parameters steer the fineness in the description of the energy dependence of electromagnetic processes in an interdependent way[3]. Various settings for the range cut and step size parameters and their influence on the neutron echo were studied in this work.

The Range Cut: By imposing a range cut through Geant4's *SetDefaultCutValue* method, one can determine the minimum range of secondary electrons, positrons and photons to be explicitly generated and tracked, while below that range, the mother particle's energy loss occurs continuously. Geant4 internally converts the range cut to a minimal required energy with respect to the properties of the medium the particle is traversing. The default value in Reference Physics Lists of the Geant4 version used in this analysis is 0.7 mm, while the absolut minimum energy is internally set to 990 eV[4].

The Database Binning: The restricted stopping power dE/dx as well as the mean free path, range and inverse range for electromagnetic processes are stored in energy-binned tables before the start of a simulation. Increasing the number of bins in these tables, i.e. the *database binning*, consequently increases the simulation accuracy. By default, the energy range from 100 eV to 10 TeV is evenly divided into 77 bins [83]. The effect of the database binning was not studied in this work but has been discussed elsewhere [82].

The Step Size: A *step* in Geant4 refers to the distance between two interactions of a given particle. The particle's mean energy loss over a step is calculated by means of the dE/dx- and inverse range-tables and fluctuations on the mean energy loss are added. Allowing for too long steps consequently decreases the accuracy of the simulation, while lowering the maximum step size will increase the required computation time. Geant4 allows the user to limit the maximum step size either directly or by setting two parameters in a dynamic *stepping function* [83]. Therein, the first parameter sets the maximum allowed value of the ratio $StepSize/Range$ and the second parameter sets the lower range threshold for the dynamic stepping. Below the lower range limit, the maximum allowed step size is equal to the particle range.

---

[3]Since these parameters only impact the modeling of electromagnetic processes, they should not affect hadronic processes. However, a dependence of the dose deposition and proton range on these parameters could for instance be observed in a proton therapy application [82].

[4]The lower energy limit corresponds to the limit of validity of the standard electromagnetic processes implemented in most Reference Physics Lists. The absolut minimum energy can, however, be adjusted by the user if low energy models for electromagnetic physics are incorporated.

Impact on the neutron echo strength and final settings

Since the parameters discussed above impact the accuracy of a simulation in a correlated way, Geant4 users are usually advised to investigate the impact on their specific application. In this analysis, the influence of a different physics list, adapted stepping function parameters and different range cuts on the neutron yield, Cherenkov yield from neutron captures and the total Cherenkov yield was studied. The results are summarized in Table 6.1: using the *QGSP_BERT_HP* list results in an overall slightly higher neutron yield. The study also showed that coarser settings in the stepping function can be compensated by sufficiently low range cut values, while the range cut parameter has almost no influence when the stepping is fine enough. The overall Cherenkov yield is not sensitive to the parameter settings.

In this analysis, the *QGSP_BERT_HP* list, together with a *DefaultCutValue* of 0.25 mm, was used[5]. All other parameters were set to their default value.

6.3   IceCube detector simulation

The modeling of high energy cascades and their delayed signal by Geant4 in the full Ice-Cube event simulation chain is computationally expensive. A routine that simulates the light yield from a neutrino interaction as well as the detector response in a time efficient way thus replaces the full detector simulation in the final analysis (see Section 7.2). For the cross-validation of this fast routine, a comparative event set from the full detector simulation was nevertheless needed.

In order to simulate the detector's light record, final state particles from the DIS process are passed to the *clsim* photon tracker [84]. Besides modeling the light yield, *clsim* is able to internally invoke Geant4 to simulate the particle steps for which Cherenkov emission is calculated. Using *clsim's* Geant4 option, the full detector simulation is therefore identical to the Geant4 stand-alone setup up to this point. Subsequently, *clsim* tracks each Cherenkov photon until it is absorbed or hits a DOM. It consequently includes the detector setup and requires information on the scattering and absorption properties of the Antarctic ice, whose depth and wavelength dependence is provided by tables characterizing the absorption and effective scattering coefficients of 10 m-thick layers. The most recent ice model, *SPICE LEA* [85], was implemented in this analysis.

In the next step, the complete detector response for photons arriving at the DOM's surface is modeled by designated IceCube simulation tools. This comprises the simulation of the PMT response including artifacts, the processing by the DOM mainboard and the involved deadtimes (see Section 3.2.1) and finally the calibration and feature extraction of the recorded waveform. The *Vuvuzela* module [71], IceCube's standard

---

[5]A range cut of 0.25 mm was used in all previous Geant4-based IceCube analyses and is thus consistently implemented here.

noise simulation routine, incorporates the modeling of correlated noise in addition to the Poissonian component (see Section 3.2.1). Applying Vuvuzela to time windows larger than around $50\,\mu$s is, however, not recommended. Vuvuzela is thus only used for the time span of the prompt signal, while Poissonian noise is manually added to the delayed signal On the uncleaned hit map, a Monopod reconstruction is run in analogy to the HESE HitSpool processing chain.

DefaultCutValue [mm]	QGSP_BERT_HP	modHadr06$_{(0.1, 100\mu m)}$	modHadr06$_{(1, 1mm)}$	QGSP_BERT_HP/modHadr06$_{(0.1, 100\mu m)}$	QGSP_BERT_HP/modHadr06$_{(1, 1mm)}$	modHadr06$_{(0.1, 100\mu m)}$/modHadr06$_{(1, 1mm)}$
Neutrons						
0.001	49.32	41.78	38.64	1.18	1.28	1.08
0.005	47.35	47.20	45.97	1.00	1.03	1.03
0.01	57.16	42.48	44.90	1.35	1.27	0.95
0.05	53.03	43.98	42.94	1.21	1.23	1.02
0.1	50.28	42.75	43.61	1.18	1.15	0.98
0.25	47.76	45.74	43.62	1.04	1.09	1.05
0.5	50.68	46.61	38.82	1.09	1.31	1.20
0.7	51.80	44.57	40.24	1.16	1.29	1.11
1	48.87	50.33	46.45	0.97	1.05	1.08
Cherenkov Photons from Neutron Captures						
0.001	6.04E+03	4.90E+03	4.60E+03	1.23	1.31	1.06
0.005	6.19E+03	5.46E+03	5.20E+03	1.13	1.19	1.05
0.01	6.66E+03	5.12E+03	5.07E+03	1.30	1.31	1.01
0.05	6.30E+03	5.44E+03	4.33E+03	1.16	1.46	1.26
0.1	5.79E+03	4.74E+03	4.35E+03	1.22	1.33	1.09
0.25	5.55E+03	5.29E+03	3.84E+03	1.05	1.44	1.38
0.5	5.82E+03	5.50E+03	3.42E+03	1.06	1.70	1.61
0.7	6.03E+03	5.08E+03	3.58E+03	1.19	1.68	1.42
1	5.58E+03	5.80E+03	3.88E+03	0.96	1.44	1.49
total Cherenkov Photons						
0.001	1.41E+06	1.49E+06	1.54E+06	0.95	0.92	0.97
0.005	1.44E+06	1.42E+06	1.39E+06	1.02	1.03	1.02
0.01	1.34E+06	1.47E+06	1.43E+06	0.91	0.94	1.03
0.05	1.36E+06	1.42E+06	1.42E+06	0.95	0.95	1.00
0.1	1.42E+06	1.41E+06	1.38E+06	1.01	1.03	1.02
0.25	1.42E+06	1.41E+06	1.39E+06	1.01	1.02	1.01
0.5	1.39E+06	1.36E+06	1.45E+06	1.02	0.96	0.94
0.7	1.39E+06	1.42E+06	1.40E+06	0.98	1.00	1.01
1	1.43E+06	1.34E+06	1.31E+06	1.07	1.09	1.02

Table 6.1: The combined influence of different physics lists, range cuts and stepping function parameters on the neutron echo strength is shown for cascades induced by $10\,\mathrm{GeV}$ π^+ particles. The tabulated numbers are the mean number per cascade of the observable stated above. Mean values are calculated from a set of 100 cascades. The subscripts on the *modHadr06* list give the parameters used in the stepping function, where $(1, 1\,\mathrm{mm})$ are the values originally implemented in Geant4's *Hadr06* example and $(0.1, 100\,\mu m)$ is the default setting. The typical uncertainty on the means is on the order of 3 % for the total Cherenkov photon count and 5 % for the count of Cherenkov photons from neutron captures and the neutron count.

7

Important ingredients of the analysis

The aim of this work is to assign every HESE cascade to one of the three major shower production mechanisms: electron neutrino charged current, tau neutrino charged current and all flavor neutral current interactions. This chapter explains the identification principle and the experimental implementation, discusses the expected performance and involved systematic uncertainties and presents the analysis strategy for individual cascades and for the ensemble of showers in the dataset.

7.1 Identification principle
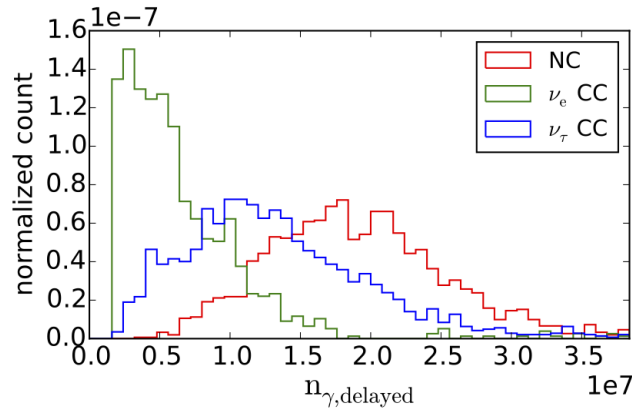
Figure 7.1: Distribution of the delayed photon count for cascades with an energy deposition between 100-150 TeV (Geant4 stand-alone simulation). Hadronic fluctuations govern the widths of the curves.

Combining the knowledge collected in Chapter 2 and Chapter 5, the three shower types are distinguishable by the strength of their neutron echo relative to the cascade energy. Only a few delayed photons are produced by neutron capture events in ν_e CC showers, while a comparatively stronger echo is generated by the multitude of neutrons in NC cascades. Showers from ν_τ CC interactions are a superposition of the hadronic shower at the interaction vertex and the electromagnetic or hadronic cascade from the tau decay and the corresponding average echo strength is thus expected in between the other shower types (see Figure 7.1). In principle, one could run the full detector simulation as

described in Section 6.3 to obtain the delayed PE distributions for any observed cascade, and attribute the shower to one of the interaction types by its measured delayed PE count. This method has the advantage of accurately incorporating detector effects for the given event hypothesis, i.e. event vertex, energy and direction, but is too CPU expensive[1] to be realized. The *Fast Simulation*, which was developed in this work to circumvent these computational limitations, is presented in the next section.

7.2 The *Fast Simulation*

The *Fast Simulation* mainly shortens the computation time required for the production of an event's delayed PE distributions by avoiding extensive Geant4 simulations. The procedure can be organized into three steps, which will be elaborated on in the following.

Step 1: the neutrino nnteraction
In this step, PYTHIA 8 (see Section 6.1) simulates neutrino interactions which lead to cascades with an energy equal to the measured energy deposition. The observable energy of an interaction is estimated from the energy of its Cherenkov light producing final state particles, taking the lower light yield of the hadronic constituents (see Figure 2.3) into account. In the ν_τ CC channel, events with a muon track from the tau decay are assumed to be classified as a track rather than a cascade and are rejected. Since the Bjorken y distribution, which determines the hadronic content and thus the echo strength in ν_e CC and ν_τ CC showers, differs for neutrinos and antineutrinos, the correct composition of particles and antiparticles is modeled. The simulation of a neutrino interaction by PYTHIA 8 is fast and was therefore not modified with respect to the full detector simulation.

Step 2: the photon count parametrization
The major ingredient in the generation of $PE_{delayed}$ distributions is the ratio of a cascade's neutron capture induced over total photon yield. Instead of running the time consuming Geant4 simulation for the reaction products from the PYTHIA 8 simulation, the *Fast Simulation* obtains the photon counts from a parametrization. The parametrization was obtained by collecting all Geant4 simulation results for a given particle type produced in the course of this work and plotting the simulated γ_{nC} and γ_{total} counts versus particle energy. The thus created histograms were smoothed using the KDE method[2]. A fit to the energy dependent mean as well as standard deviation of the photon yield now allows for drawing γ_{nC} and γ_{total} counts for any so far simulated particle type at any desired energy. Figure 7.2 exemplarily shows the smoothed histograms together with the fit results for positively charged

---

[1]Roughly 250,000 CPU hours are needed to simulate 200 events for each of the three interaction types for a 100 TeV shower.

[2]KDE (Kernel Density Estimation) is a non-parametric routine to estimate the underlying probability density function of a random variable given a measured outcome. KDE is especially useful in low statistics problems. Here, a Gaussian kernel was used.

kaons. The corresponding plots for the remaining particle types can be found in Appendix B.
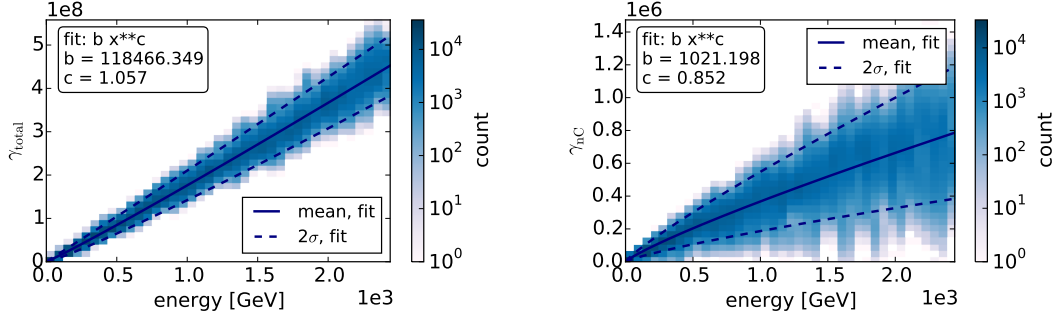
Figure 7.2: The KDE-smoothed histograms of the respective photon count versus energy of the K$^+$ primary. The mean and standard deviation is calculated in each energy bin and is fitted as a function of the form $b \times (\text{energy}/\text{GeV})^c$.

Step 3: detector effects and resolution

The true ratio of delayed over prompt photons is altered by detection artifacts (see Section 3.2.1) as well as the limited resolution due to the delayed photon statistics. While the limited resolution of the detector and the likelihood method mainly broadens the distributions, detector effects have a more complex impact. The three most important detection artifacts are:

saturation: A significant fraction of the prompt signal is not recorded when the PMTs reach their saturation region in high energy showers.

directionality: While the neutron capture signal spreads isotropically from the event vertex, Figure 7.4 illustates that, due to the Cherenkov cone, the detectable prompt PE count strongly depends on the event direction.

DOM deadtime: During deadtime periods, DOMs are unavailable for the recording of the delayed signal (see Section 3.2.1). As a consequence, the exponential shape of the delayed signal is distorted as exemplarily shown in Figure 7.3, and a substantial portion of up to 40% of the delayed photons is not observable in the detector. In the approximately first $10\,\mu$s, in which the prompt signal is recorded, deadtimes play a negligible role.

signal detection efficiency: Compared to an event's prompt light, the neutron echo is a very faint signal. Not every delayed photon hitting a DOM thus leads to a detectable pulse.
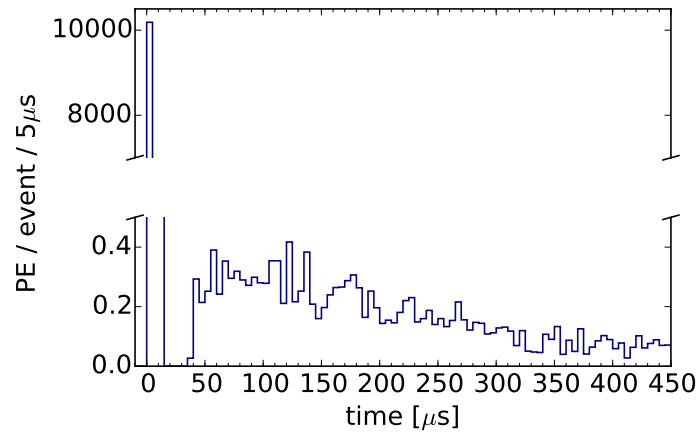
Figure 7.3: The stacked PE time spectrum in the five brightest DOMs of around 100 NC showers with a reconstructed energy between 100-150 TeV, as simulated by the full detector simulation. DOM deadtimes cause a clear gap between the prompt and delayed signal.
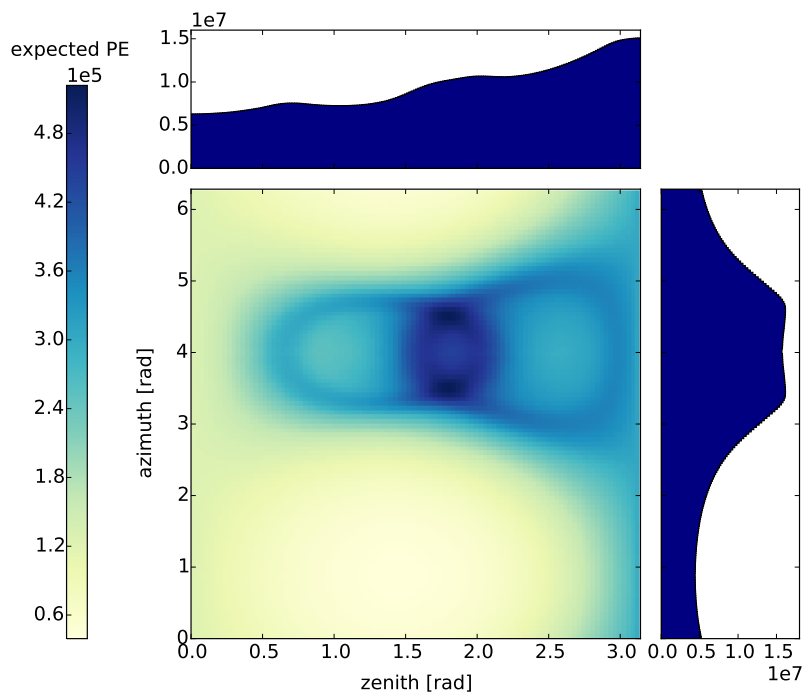
Figure 7.4: The dependence of the prompt PE expectation in the five brightest DOMs on the event direction for a 150 TeV cascade. The effect of the Cherenkov cone almost directly hitting the modules nicely reflects in an azimuthal dependence (dark blue ring).

The simulation of the delayed PE count expected in the experiment is based on the assumption that the measured ratio of delayed over prompt Cherenkov photons, $\frac{n_m^d}{n_m^p}$, equals the ratio obtained from the Geant4 parametrization, $\frac{n_G^d}{n_G^p}$, times factors correcting for the detector effects discussed before:

$$\frac{n_m^d}{n_m^p} \quad = \quad \frac{n_G^d}{n_G^p} \times c_{sat} \times c_{dir} \times c_{dt} \times c_{\eta} \qquad . \tag{7.1}$$

The calculation of the correction factors crucially relies on the photon yield prediction from the *photontables*, which were briefly introduced within the context of the event reconstruction in Section 5.2. The *photontables* return the PE expectation in every DOM given an event hypothesis as input and already include all detection efficiencies such as the module's angular acceptance and quantum efficiency.

The saturation correction factor c_{sat} equals the ratio of the *photontables*' saturation-free PE prediction for the given event reconstruction over the measured prompt PE count:

$$c_{sat} = \frac{PE_{reco}^{photontables}}{PE_{\,measured}} = \frac{PE_{reco}^{photontables}}{n_m^p} \qquad .$$

The directionality correction factor c_{dir} is given by the photontable prediction averaged over event directions divided by the PE count returned by the *photontables* for the particular event reconstruction:

$$c_{dir} = \frac{\left\langle PE_{reco}^{photontables} \right\rangle |_{directions}}{PE_{reco}^{photontables}} \qquad .$$

The deadtime correction factor c_{dt} is calculated by a fast algorithm developed within the scope of a Bachelor's thesis [86, 87], since the use of the full detector simulation to model the time dependent deadtime correction with sufficient statistics is computationally inefficient. As the concrete deadtime pattern depends on the coincidence conditions imposed by the prompt light, afterpulses and the delayed light itself (see Section 3.2.1), the modeling of the deadtime effects is complex. The input-file for the Python-based routine holds the prompt PE expectation at every module as well as basic event information such as vertex, energy and event direction. Such files can be individually produced for every cascade by means of the *photontables*. The algorithm then approximates the smearing of the prompt signal due to photon scattering in the ice by the Pandel function [88] and models afterpulses according to laboratory measurements of the IceCube PMTs. The delayed signal is assumed to be 0.1 % of the prompt PE expectation at every DOM and is temporally distributed following an exponential decay with a decay time of 217 μs. All modeled light signals are Poisson-randomized. Based on the simulated signals,

a coincidence matrix stores the HLC and SLC condition for the modules. Given the recorded photons and coincidence information, the program subsequently models the behavior of the ATWDs and FADCs during waveform capture and digitization in analogy to the full detector simulation. As a simplification, the characteristic curve of the PMTs is assumed to be linear. Lastly, a deadtime correction function for the delayed signal is determined and the program returns four lists holding the HLC and SLC hit times with and without the deadtime mask applied. In order to produce smooth time distributions, the algorithm can superimpose the results from many Toy Monte Carlo trials. The deadtime correction factor c_{dt} is then given by the ratio of the number of delayed hits including deadtime effects over the deadtime-free delayed hit count (see Figure 7.5).
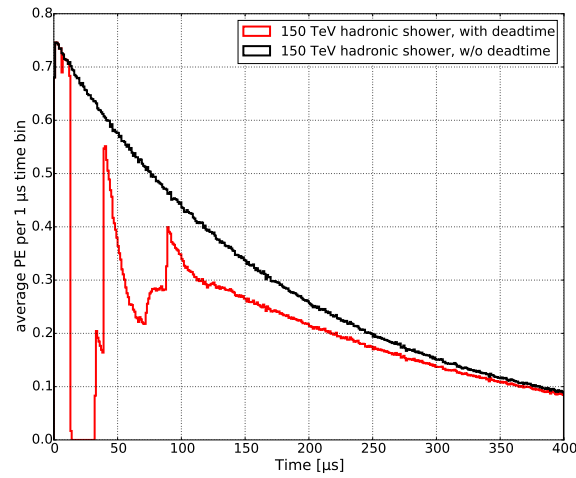
Figure 7.5: The deadtime-free neutron echo (black curve) and the neutron echo with the deadtime mask applied (red curve) in a 150 TeV hadronic shower, as simulated by the deadtime correction algorithm [86]. A signal decay time of $190\,\mu$s was assumed and SLC and HLC hits were combined for this plot.

The efficiency correction factor c_η was determined by comparing the number of delayed photons arriving at the DOMs to the resulting output pulses in Figure 7.6. A fraction of 0.81 of the delayed photons is registered by the modules.
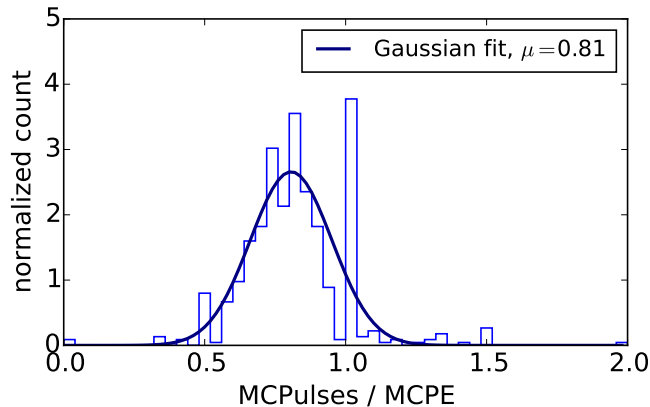
Figure 7.6: The number of delayed Monte Carlo simulated pulses (MCPulses) over the number of Monte Carlo delayed PE (MCPE) for around 600 showers at $100 - 150\,\mathrm{TeV}$ energy. The peak at 1 is caused by the integer nature of the pulse and MCPE count.

With the terms derived for c_{sat} and c_{dir} inserted into Equation 7.1, the expectation for the measured number of delayed photons reads:

$$n_m^d \quad = \quad \frac{n_G^d}{n_G^p} \times \langle PE_{reco}^{photontables}\rangle|_{directions} \times c_{dt} \times c_\eta \qquad . \qquad (7.2)$$

Multiple n_m^d obtained by Equation 7.2 result in delayed photon count probability densitiy distributions as shown in Figure 7.1. To properly account for hadronic fluctuations, PYTHIA 8 simulates 200 neutrino interactions, which would deposit the measured cascade energy, for each interaction type. Afterwards, photon counts are drawn for the reaction products in Step 3 of the *Fast Simulation*. After the correction factors have been applied, the final counts are smeared to incorporate the signal hit dependent resolution (see Figure 7.12).

7.2.1 Cross-validation of the *Fast Simulation*

Though extremely CPU intensive, the full detector simulation was run for around 200 events in each interaction channel to test whether the detector response is correctly modeled in the *Fast Simulation*. The events were all simulated with an energy deposition of $100 - 150\,\mathrm{TeV}$ and at the same vertex[3], but with scrambled event directions.

---

[3]The vertex of one of IceCube's PeV showers ($x \approx -77\,\mathrm{m}$, $y \approx 266\,\mathrm{m}$, $z \approx 25\,\mathrm{m}$, with respect to the center of the detector) [42] was chosen as a reference point at the start of this work. This choice was continued in the full detector simulation throughout the thesis.

As for data, the delayed light signal is extracted from the simulated pulsemaps using a likelihood approach (see Section 7.4). This measured $PE_{delayed}$ count can then be compared to the prediction for the shower's known particle content obtained in step 2 and step 3 of the *Fast Simulation*, as shown for one sample event in Figure 7.7. On
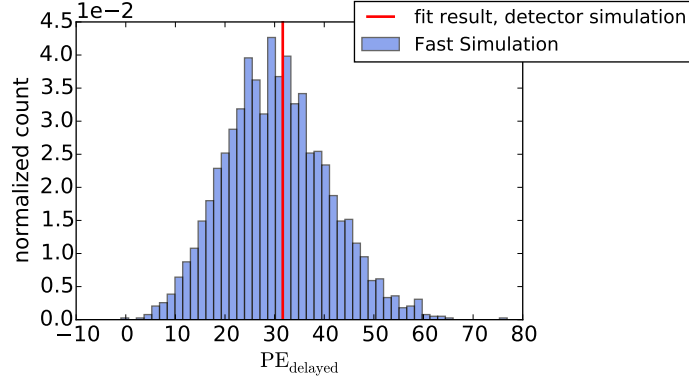
Figure 7.7: *Fast Simulation* results and fitted $PE_{delayed}$ count from a full detector simulation for a NC cascade.

average, the mean of the *Fast Simulation* distribution and the fitted $PE_{delayed}$ count should be the same such that the ratio of both quantities equals one. When taking the natural logarithm of the ratios, a Gaussian distribution with mean zero is thus expected. The corresponding distributions in Figure 7.8 demonstrate that the *Fast Simulation* is consistent with the full detector simulation at the 10% level. The combined *Fast Sim-*

(a) ν_e CC   (b) ν_τ CC   (c) NC
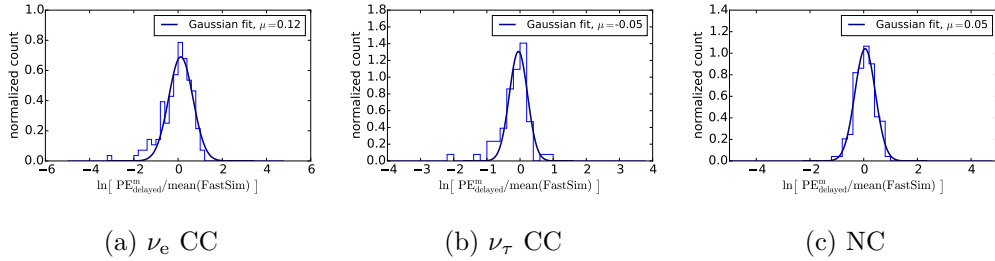
Figure 7.8: The *Fast Simulation* is run for each event's true particle content and the thus created $PE_{delayed}$ distribution is compared to the event's fitted $PE_{delayed}$ value. On average, the mean of the distribution from the *Fast Simulation* should equal the measured $PE_{delayed}$ count.

ulation and full detector simulation results in Figure 7.9 additionally allow for a visual comparison.

|(a) ν_e CC | (b) ν_τ CC | (c) NC |
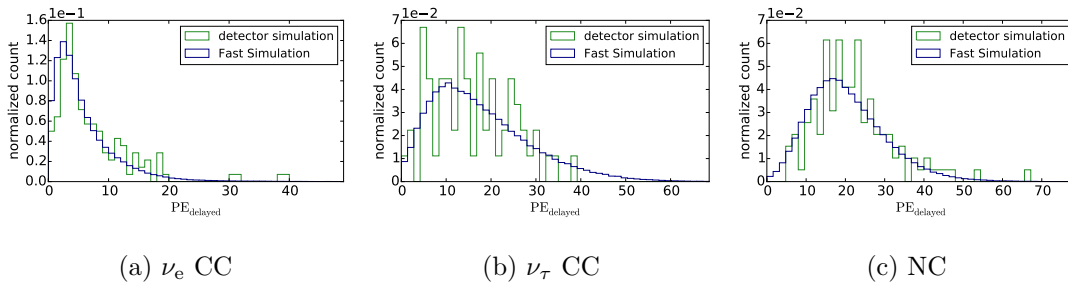
Figure 7.9: The plots show the roughly 200 fitted $PE_{delayed}$ values in each interaction channel in green compared to the combined $PE_{delayed}$ distributions from the *Fast Simulation* in blue for well reconstructed events.

7.3 Likelihood ratio test, p-value and sensitivity

A method for the interpretation of results from a likelihood analysis is needed in the evaluation of the echo fitting routine (see next section) as well as in the final analysis of the data sample (see Section 7.7 and Chapter 8). This section briefly summarizes the fundamental notions of a widely used approach.

A convenient tool for a statistical interpretation of likelihood results is the study of a *test statistic* (TS). When testing which of two proposed hypotheses describes data best, the test statistic is commonly defined as the ratio of the maximum likelihood under the null hypothesis H_0 and the maximum likelihood under the alternative hypothesis H_i:

$$\text{TS} = -2 \cdot \ln \left(\frac{\mathscr{L}_{H_0}}{\mathscr{L}_{H_i}} \right) \qquad . \tag{7.3}$$

Equation 7.3 shows that H_i is favored over the null hypothesis for larger values of TS.

The p-value of an outcome is defined as the percentage of Toy Monte Carlo experiments simulated under the null hypothesis that produce a higher test statistic value than the outcome itself. The p-value thus states the probability that the observed outcome is the results of a pure fluctuation of the null hypothesis model.

Investigating the energy threshold from which the likelihood method presented in the next section is able to discriminate the neutron echo from pure noise, H_0 is the noise-only hypothesis and H_i is the signal hypothesis assuming the presence of an echo. For such a setup, the sensitivity is defined as the signal strength which produces a p-value smaller than 0.5 in 90% of the runs. In the following section, the signal strength is represented by the hadronic shower energy yielding a certain echo strength.

7.4 Signal extraction, resolution and energy threshold

A likelihood approach is used to extract the number of neutron capture induced signal PE, n_s, from an event's pulsemap. With N being the total number of delayed PE in the analysis time window $[t_1, t_2]$, the likelihood function is given by[4]:

$$\mathscr{L}(N \,|\, n_s) \;=\; \frac{e^{-\mu_b} \cdot \mu_b^{(N-n_s)}}{\Gamma(N - n_s + 1)} \prod_{i=1}^{N_{DOM}} \prod_{j=1}^{N_{hit}} \left(\frac{n_s}{N} \cdot P_s(\vec{r}_i, t_j) + \frac{N - n_s}{N} \cdot P_b(t_j) \right) \quad . \tag{7.4}$$

The signal probability density function P_s decomposes in a spatial and temporal part. The latter includes the exponential shape of the signal time spectrum and the time dependent deadtime correction factor $f_{dt}(t_j)$ for hit j in DOM i:

$$P_s(\vec{r}_i, t_j) \;=\; P_s(\vec{r}_i) \times P_s(t_j) \tag{7.5}$$

$$= \left(\frac{q_i}{\sum_{i=1}^{N_{DOM}} q_i} \right)\Bigg|_{\text{directional mean}} \times \; \frac{f_{dt}(t_j) \cdot e^{-t_j/\tau}}{\int_{t_1}^{t_2} f_{dt}(t) \cdot e^{-t/\tau} dt} \quad . \tag{7.6}$$

$f_{dt}(t)$ is calculated by the algorithm introduced in Section 7.2 and is given by the fraction of expected hits with deadtime effects taken into account over the deadtime-free hit expectation. q_i is the charge expectation in DOM i.

In addition to the uniform noise assumption of $\mu'_b \approx 560 \, \text{Hz}$ per DOM[5], the background probability density function P_b also includes the deadtime correction:

$$P_b(t_j) \;=\; \frac{1}{N_{DOM}} \cdot \frac{f_{dt}(t_j)}{\int_{t_1}^{t_2} f_{dt}(t) dt} \quad . \tag{7.7}$$

A pre-factor in the likelihood function allows for Poissonian fluctuations around the mean noise expectation of $\mu_b = \mu'_b \cdot N_{DOM} \cdot (t_2 - t_1) \cdot \frac{\int_{t_1}^{t_2} f_{dt}(t) dt}{(t_2 - t_1)}$.

The sensitivity, in terms of a minimal hadronic shower energy, and resolution achievable by the likelihood approach is determined by means of Toy Monte Carlo experiments, in which the expected signal and noise for different event hypotheses is modeled. The likelihood method is then assessed based on its performance on the simulated data, using the tools presented in Section 7.3. Early stages of this algorithm can also be found in [55].

Since modules too distant from the event vertex have a low probability of detecting the neutron echo and thus mostly contribute noise hits, the question of the ideal number of DOMs[6] to be included in the analysis arises. While the resolution should in

---

[4]Note that the factorial $(N - n_s)!$ has been approximated by $\Gamma(N - n_s + 1)$ to allow for non-integer values of n_s.

[5]The noise rate is individually calculated for an event's contributing DOMs from the 0.5 s of HitSpool data before the HESE trigger.

[6]The importance of a DOM is ranked by its prompt light expectation obtained from the *photontables*.

principle improve with the number of DOMs used in the fit, systematic uncertainties due to the modeling of noise and detector artifacts would rise. Figure 7.10 and Figure 7.11 show the relative resolution and achievable sensitivity in dependence of the number of included DOMs, respectively. The precise number of modules, in which the signal is searched for, has negligible impact on the sensitivity. As eight DOMs have been considered in the analysis in early stages of this work, this choice was maintained. The corresponding relative resolution in dependence of the mean fitted signal hits is shown in Figure 7.12.

Figure 7.11 in addition suggests that, for the simulated scenario, the analysis is able to distinguish the neutron echo from pure noise from around 21 TeV hadronic cascade energy on. It must, however, be noted that a fairly optimistic assumption for the echo strength has been made in these studies: the signal was modeled as a fixed fraction (0.07 %) of the prompt light at every module and an event direction with a high prompt light yield (see Figure 7.4) was assumed. Furthermore, all events were simulated at a vertex which is only 20 m distant from the DOM with the highest light expectation. In order to explore a pessimistic scenario, the sensitivity for an event direction with the fewest prompt light expectation and a vertex 50 m away from the brightest module was determined to be around 190 TeV. The sensitivity threshold for average showers thus lies around 100 TeV hadronic energy.
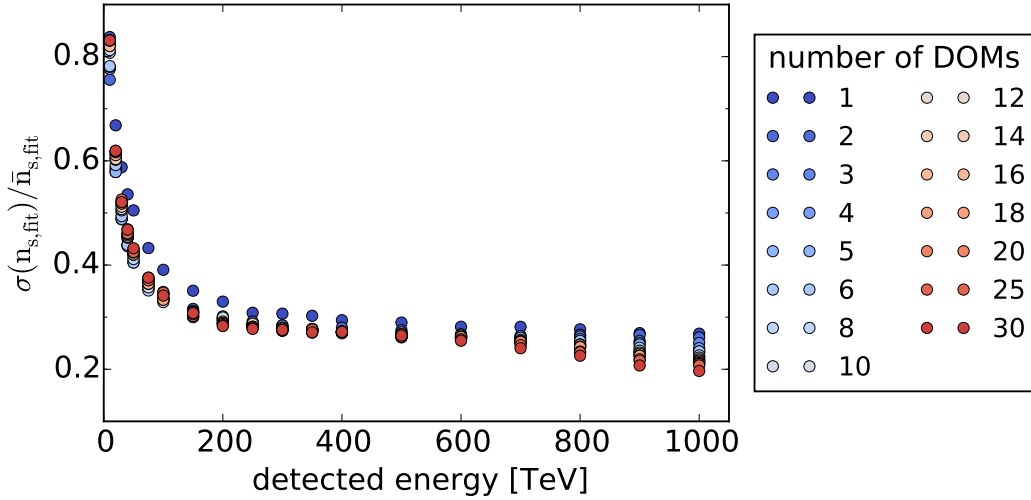
Figure 7.10: The relative resolution versus hadronic shower energy for different numbers of DOMs included in the analysis. The event vertex was fixed at $x = -77\,\mathrm{m}$, $y = 266\,\mathrm{m}$, $z = 25\,\mathrm{m}$ in this study (see Section 7.2.1).
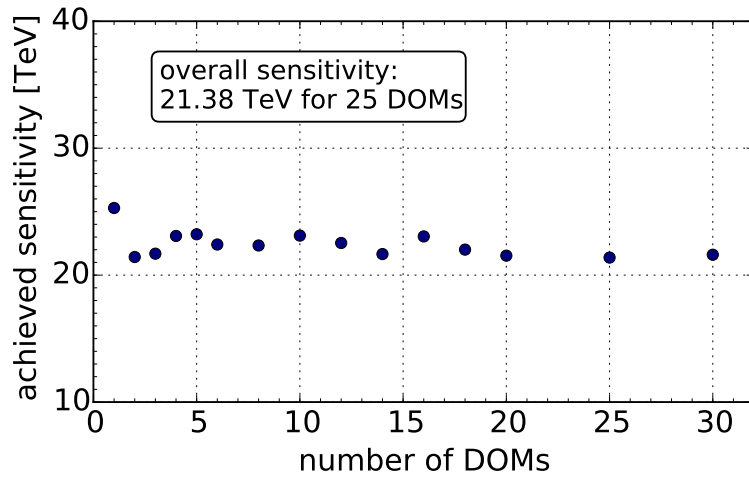
Figure 7.11: The achievable sensitivity versus the number of included modules for a shower vertex 20 m distant to the nearest DOM (optimistic scenario). The event vertex was fixed at $x = -77$ m, $y = 266$ m, $z = 25$ m in this study (see Section 7.2.1).
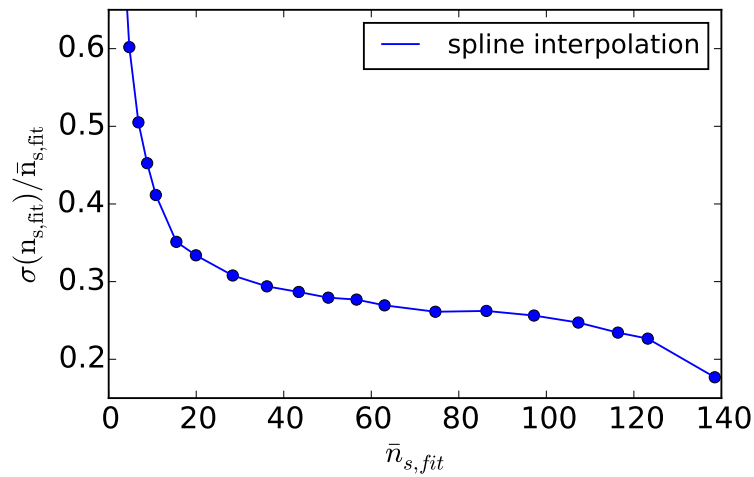
Figure 7.12: The relative resolution versus the mean fitted signal hits with eight DOMs considered in the analysis. The event vertex was fixed at $x = -77$ m, $y = 266$ m, $z = 25$ m in this study (see Section 7.2.1).

In order to study the extent to which correlated noise affects the fit performance, a *Vuvuzela*-like module (see Section 6.3), modeling the correlated noise in addition to the Poissonian component, was implemented in the Toy Monte Carlo. An average noise model was assumed for all DOMs with parameters taken from [71]. Figure 7.13 shows the impact of correlated noise on the relative resolution compared to the uniform-only noise assumption. Taking correlated noise into account, the relative resolution decreases by five percent points at 75 TeV hadronic energy, while the difference becomes negligible with increasing shower energy.

The spread in $\sigma(n_{s,fit})$, reflected by the relative resolution, includes the statistical as well as the fit uncertainty. To assess the fit uncertainties individually, Figure 7.14 displays the mean and standard deviation of the $(n_{s,fit} - n_{s,true})/n_{s,true}$ distribution at different hadronic shower energies. With the mean values being close to zero (less than 5 % deviation at 75 TeV), this also verifies that the number of injected delayed photons is, on average, reproduced by the fit in both noise scenarios. The spread in the fitted signal hits is larger when correlated noise is included in the simulation. This is, however, a subdominant effect as Figure 7.13 demonstrates that this spread is dominated by statistical fluctuations.
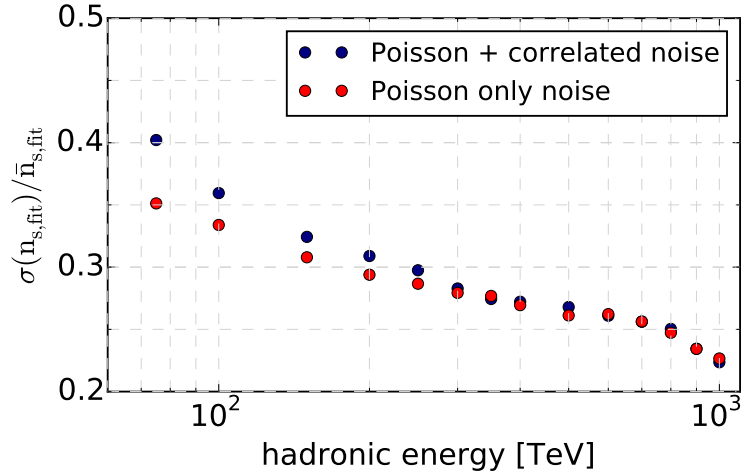
Figure 7.13: The relative resolution versus hadronic shower energy for a *Vuvuzela*-like and Poisson-only noise hypothesis. Eight DOMs were included in the analysis and the event vertex was fixed at $x = -77$ m, $y = 266$ m, $z = 25$ m in this study (see Section 7.2.1).
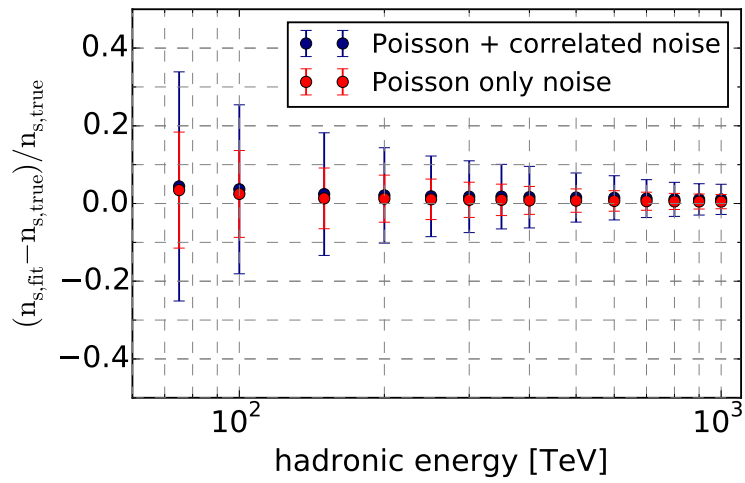
Figure 7.14: The fit uncertainty with and without correlated noise taken into account is shown in blue and red, respectively. The marker represents the mean of the $(n_{s,fit} - n_{s,true})/n_{s,true}$ distribution of multiple Toy Monte Carlo runs at a given energy and the error bar refers to the distribution's standard deviation, not the error on the mean. The event vertex was fixed at $x = -77\,\mathrm{m}$, $y = 266\,\mathrm{m}$, $z = 25\,\mathrm{m}$ in this study (see Section 7.2.1).

7.5 Systematic uncertainties

Since many uncertainties cancel in the ratio of delayed over prompt Cherenkov signals, its error is dominated by the uncertainty on the overall neutron yield in particle showers and on the neutron capture occurrence. These numbers are, however, subject to large hadronic and nuclear uncertainties in any event simulation.

7.5.1 Hadronic and nuclear uncertainties

Hadronic and nuclear interactions are associated with large fluctuations and are less explored in experiments than the electromagnetic sector. The modeling of hadronic processes is a complex topic and the involved uncertainties are difficult to assess. The influence of this class of uncertainties on an echo analysis in IceCube was estimated by a comparison of isotope yields from cosmic-muon-spallation measured in low energy neutrino detectors like *Borexino* [89], *KamLAND* [90] and *Super-Kamiokande* [91] and simulation results from Geant4 and *FLUKA*[7] [25]. These studies found experiment and simulation to agree roughly on a factor 2 level.

Similar uncertainties were found in a study of the fluctuation in the neutron yield per 250 MeV proton interacting in water due to the hadronic processes implemented in the Geant4 simulation [92]. Further work investigated the variation in the neutron production due to the different hadron inelastic scattering models in the Reference Physics Lists *QGSP_BERT_HP* and *QGSP_BIC_HP* [93] (see Appendix A for details on the physics lists). Therein, simulation was compared to data from $\mathcal{O}(\text{GeV})$ protons interacting in high purity lead. A factor of about 1.5 difference between the results from both physics lists can be read off in Figure 7.15.

Combining these results, the influence of hadronic and nuclear uncertainties in the event simulation can be quantified by a typical factor of 2. Assuming a symmetric Gaussian error, this would correspond to a 33 % uncertainty. However, when comparing the neutron capture occurrence simulated with the Geant4 version and settings used in this work to results from FLUKA [25] in Figure 7.16, an uncertainty of 45 % on the Geant4 prediction is required to cover the FLUKA simulation completely.

---

[7]As an alternative to Geant4, FLUKA is another simulation package commonly used in particle physics.
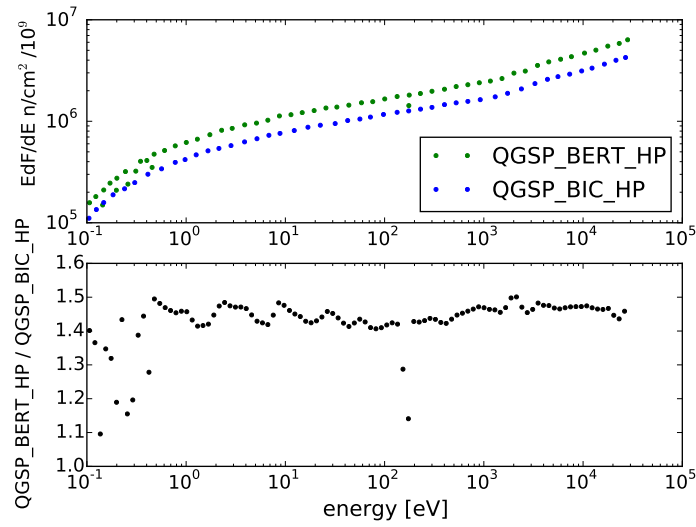
Figure 7.15: The most relevant part of Fig. 5. in [93]. From the plot, showing the neutron fluence versus neutron energy measured by the *TARC* experiment, a factor 1.5 difference between the two Reference Physics Lists *QGSP_BERT_HP* and *QGSP_BIC_HP* can be read off. The experimental data lie between the two simulated fluences [93].
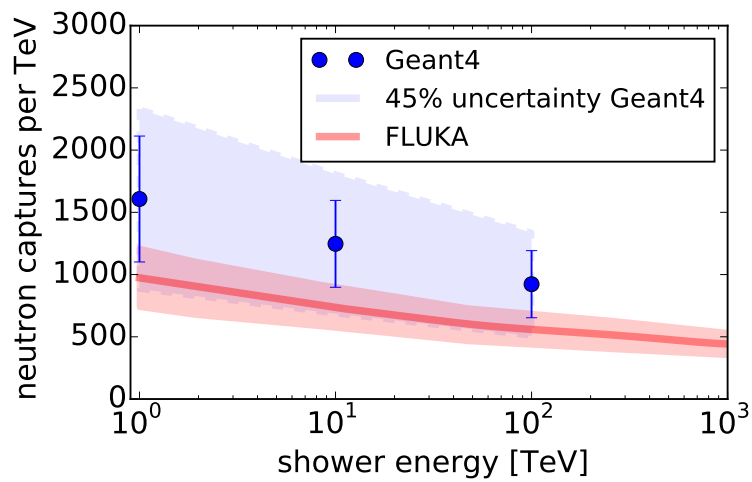
Figure 7.16: The neutron capture occurrence in FLUKA [25] and Geant4. The red band around the FLUKA mean and the error bars on the Geant4 simulation show the 1σ intrinsic fluctuations.

7.5.2 Photon count parametrization

As discussed in Section 7.2, photon counts entering Equation 7.2 are obtained from a parametrization of Geant4 simulations. The uncertainty involved with the parametrization itself and the extrapolation towards higher particle energies is estimated through a comparison of the *Fast Simulation* and photon counts obtained from IceCube's photon propagation package *clsim*, which internally invokes Geant4 for particle propagation and subsequently models the Cherenkov yield in the full detector simulation (see Section 6.3). Since the high energy electron determines the echo strength in electromagnetic showers and pions are the dominant momentum carriers in hadronic cascades, the comparison focuses on these particle species. The *clsim* module was thus initiated with e^- and π^+ primaries at energies between $500\,\mathrm{GeV}$ and $100\,\mathrm{TeV}$ and compared to the corresponding *Fast Simulation* prediction[8] in Figure 7.17 and Figure 7.18 [9]

The reference point for the uncertainty estimation is $\mathcal{O}(100\,\mathrm{TeV})$ shower energy. For electromagnetic cascades, the *Fast Simulation* and *clsim* are therefore compared at $100\,\mathrm{TeV}$ electron energy. To obtain the typical momentum of hadronic particles at these shower energies, the corresponding energy distributions were studied for 100-150 TeV hadronic cascades (see Appendix C). The highest value for the ninetieth percentile can be found in the $\mathrm{K}_{\mathrm{L}}^0$ energy distribution at $14\,\mathrm{TeV}$, which serves as a conservative reference point. The deviation in the relative echo strength from *clsim* and the *Fast Simulation* at the reference energy is around $3\,\%$ and $20\,\%$ for electrons and pions, respectively. A $20\,\%$ uncertainty is thus conservatively assumed on the photon count parametrization.

(a) electron, γ_{total} (b) electron, γ_{nC}
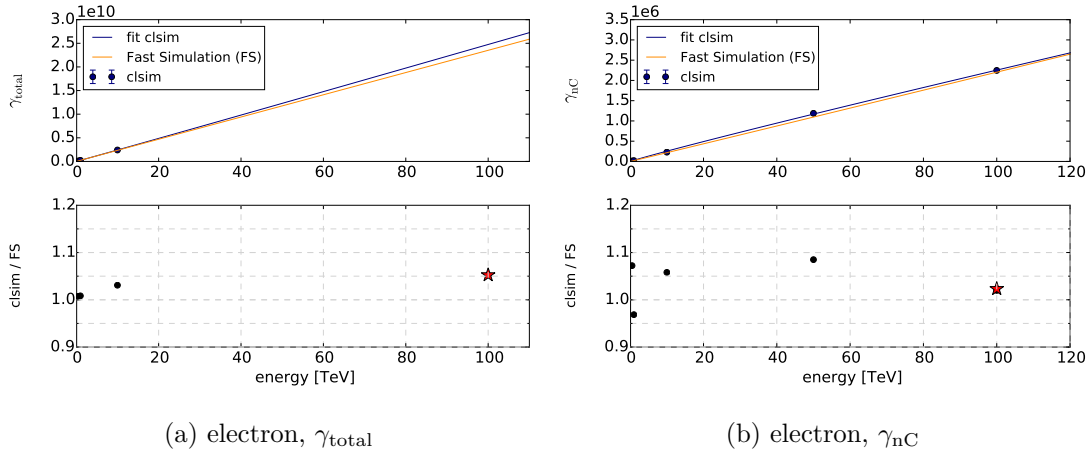
Figure 7.17: Photon counts from the *Fast Simulation* compared to *clsim*. The red star marks the ratio of the fitted *clsim* result over the *Fast Simulation* at $100\,\mathrm{TeV}$.

---

[8]For hadronic showers, a realistic particle composition was assumed in the *Fast Simulation*, since possible deviations in the parametrization are expected to be attenuated when considering the mixture of an interaction's final state particles.

[9]Above 50 TeV, the counter for the total number of Cherenkov photons implemented in *clsim* overflowed. The results are thus not included in Figure 7.17 (a) and Figure 7.18 (a).

(a) pion, γ_{total} (b) pion, γ_{nC}
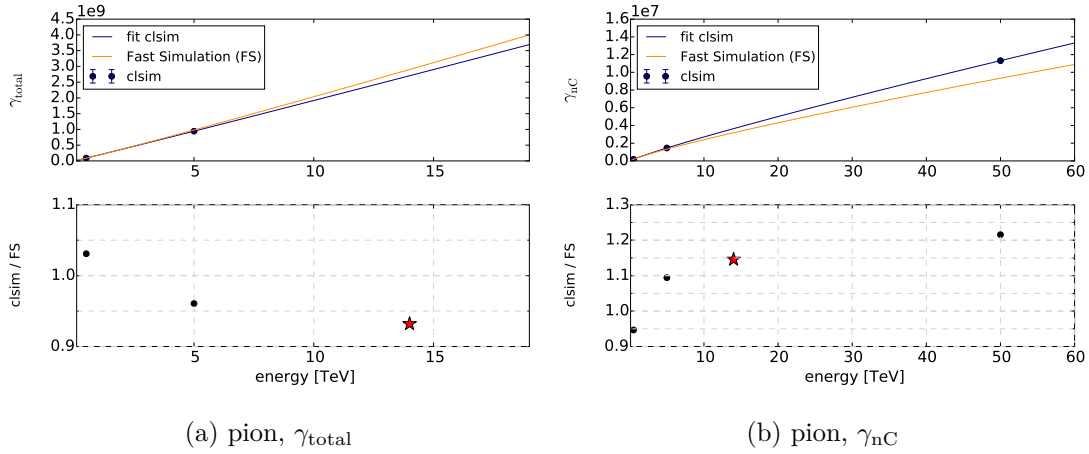
Figure 7.18: Photon counts from the *Fast Simulation* compared to *clsim*. *clsim* was initiated with positively charged pions, while the *Fast Simulation* result was obtained for a realistic mixture of hadronic particles. The red star marks the ratio of the fitted *clsim* result over the *Fast Simulation* at 14 TeV.

7.5.3 DOM deadtime calculation

The deadtime correction factor is most sensitive to the prompt signal strength and the corresponding afterpulses [86]. The prompt signal influences the number of ATWD channels to be read out while strong afterpulses, whose magnitude varies from DOM to DOM, can cause local coincidences, thus impacting the required digitization time. The prompt signal strength is determined by means of the *photontables* and the involved uncertainty will be discussed is the next section.

The algorithm assumes a fixed amount of 0.1 % of the prompt photon count at each DOM as the expectation value for the delayed signal. In contrast to the directional dependence of the prompt light yield, however, the neutron capture signal spreads isotropically. Imposing the same directionality on the delayed signal can thus falsely estimate the true echo strength at a given module. Studies showed that varying the signal assumption by a factor 3, thereby reflecting the photon yield range in dependence of the event direction in Figure 7.4, can scale c_{dt} by up to 20 %. Although the uncertainty involved in the calculation of the deadtime correction factor is probably smaller, an error of 20 % is assumed in this work, thus conservatively covering additional approximations made in the simulation of the photon scattering and the PMT response (see Section 7.2).

7.5.4 Noise simulation

While correlated noise influences the statistical power of the likelihood fitting method (see Section 7.4), the effect is small for the considered shower energies and minor compared to the dominating effect of hadronic fluctuations. The uncertainty only indirectly enters the analysis in case that a sequence of correlated noise hits mimics the delayed neutron signal, as discussed in Section 5.1.2.

7.5.5 Event misreconstruction

The event reconstruction is introduced into the analysis method via the deadtime correction factor and the averaged PE count in Equation 7.2. Both quantities are calculated based on the photon yield prediction taking the reconstruction as event hypothesis.

The expected signal loss due to DOM deadtimes depends on the position of the shower vertex relative to the DOMs, the event direction and the deposited energy: the higher the number of photons arriving at the DOMs, the larger are the expected deadtime effects. The impact of the limited resolution in the event reconstruction on the deadtime correction factor was studied by means of 198 NC showers in IceCube, all simulated in the full detector simulation chain with the same interaction vertex and direction and a deposited energy in the range of $100 - 150\,\text{TeV}$. A Monopod reconstruction, using minuit's MIGRAD [94] algorithm to look for the minimum in the negative log-likelihood space, is run on the simulated events. This must be contrasted with the Monopod reconstruction for actual HESE showers, in which the minimum is found by scanning the entire log-likelihood space. This *brute force* method does not run the risk of getting stuck in a local minimum, thus achieving an overall better resolution. The angular resolution for cascades above $100\,\text{TeV}$ is on the order of $15\,°$ [34], while $10\,\text{m}$ is a conservative estimate for the vertex resolution [95].

Figure 7.19 shows the dependence of the deadtime correction factor on the reconstruction parameters (subfigures (a), (c) and (e)) as well as its variation within the expected resolution (subfigures (b) and (d)) . While a negligible dependence on the reconstructed energy is observed within the resolution range of around $15\,\%$ for high energy cascades [44], the combined uncertainty of the deadtime correction factor due to limited vertex and directional resolution is around $7\,\%$.

The events, based on which the uncertainty on the deadtime correction factor was determined, represent an extreme case, as the distance between the vertex and the closest string is small (approximately $20\,\text{m}$). Since the correction factor becomes less sensitive to variations in the string-to-vertex-distance the further away vertices are from the closest string [86], the error estimation is conservative.

Because the PE count gets averaged over event directions, its uncertainty is not sensitive to a possible directional misreconstruction. The $15\,\%$ energy resolution, however, translates directly into an uncertainty of the same amount on $\langle\text{PE}_{\text{reco}}^{\text{photontables}}\rangle|_{\text{directions}}$, as the photon yield scales approximately linearly with energy. The $15\,\%$ error already includes

further relevant systematic uncertainties such as vertex misreconstruction, the efficiency of DOMs in transferring incident light into an electrical signal and the modeling of ice properties [44].

7.5.6 Summary of systematic uncertainties

Table 7.1 summarizes the systematic uncertainties assumed in this analysis. A combined error of $\pm$ 56 % is associated with the expected number of delayed PE, which is dominated by the large uncertainty in the hadronic and nuclear physics modeling. Most of the uncertainty estimates are rather conservative. However, given that this is the first analysis of this kind, such an approach is considered appropriate.

systematic effect	uncertainty
hadronic and nuclear physics	45 %
photon count parametrization	20 %
c_{dt}: calculation method	20 %
c_{dt}: event misreconstruction	7 %
averaged PE: energy resolution, including other systematic uncertainties	15 %
combined uncertainty	**$\pm$ 56 %**

Table 7.1: Summary of systematic uncertainties
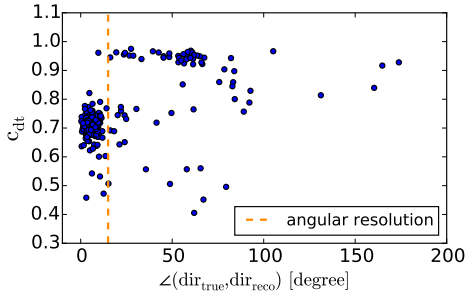
(a) correlation between c_{dt} and the angle between true and reconstructed direction. The dashed orange line marks the assumed angular resolution of 15 degree.
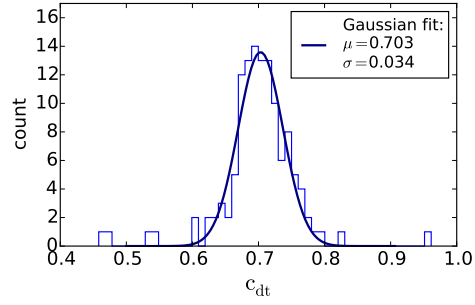
(b) c_{dt} distribution for events with $\angle(dir_{true}, dir_{reco}) \leq 15$ degree.
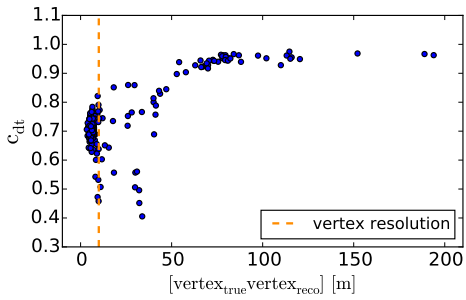
(c) correlation between c_{dt} and the distance between true and reconstructed vertex. The dashed orange line marks the assumed vertex resolution of 10 m.
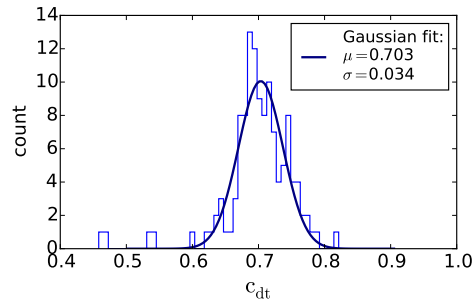
(d) c_{dt} distribution for events with $[vertex_{true}, vertex_{reco}] \leq 10$ m.
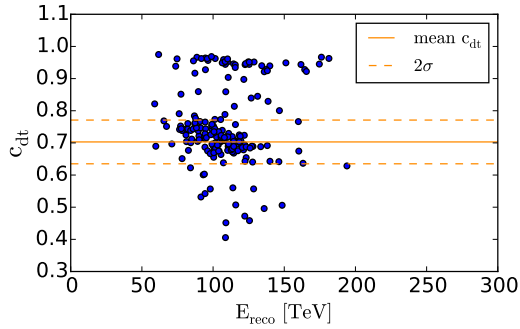
(e) correlation between c_{dt} and the reconstructed energy. The solid and dashed orange lines mark the mean and 2σ range of the c_{dt} distributions in subfigure (b) and (d).

Figure 7.19: The uncertainty on the deadtime correction factor due to event misreconstruction within the uncertainties expected for high energy cascades is around 7 % . The correlation between c_{dt} and the reconstructed energy is a subdominant effect. The fitted mean and standard deviation of the c_{dt} distributions in figure (b) and (d) are identical within the given digits, since the selection of well reconstructed events causes a large overlap of the event subsamples.

7.6 The HESE simulation dataset

A high-statistics simulation dataset, produced within the IceCube collaboration for HESE-related studies, is used in this work to validate the *Fast Simulation* for a real-case scenario, to calculate the expected event rates for the 22 month livetime of the analysis and to demonstrate the analysis approach for the ensemble of events on a realistic Monte Carlo sample. As the high-statistics HESE simulation dataset is primarily intended to be used in standard IceCube analyses, it does not include the CPU-intensive echo simulation.

7.6.1 Full detector simulation for 13 representative events

In order to confirm that the *Fast Simulation* correctly predicts the $PE_{delayed}$ distributions for a realistic dataset, the delayed light yield simulation by the full detector simulation chain was added to 13 events from the high-statistics HESE simulation dataset. Around 13 events can be expected in the 22 months livetime of this analysis (see Section 7.6.2) and energy, vertex, event direction and interaction type composition of the selected events correspond to a real-case scenario.

For each event, a compilation of three plots was produced, as illustrated in Figure 7.20 for a sample ν_τ CC shower: at the top, an event view of the simulated cascade is shown. The histogram in the lower left corner displays the $PE_{delayed}$ distribution generated by the *Fast Simulation* for the cascade's known particle content. Therein, the best-fit $PE_{delayed}$ count for the full detector simulation is marked red. On the bottom right, the $PE_{delayed}$ distributions, simulated for an unknown shower content (see Section 7.2), are shown together with the best-fit $PE_{delayed}$ value. Plots for the remaining events can be found in Appendix D.

To assess whether the *Fast Simulation* and the full detector simulation yield consistent results, a comparison of the $PE_{delayed}$ distribution, produced with the shower's true particle content, and the full detector simulation's best-fit $PE_{delayed}$ value is most meaningful. For all 13 events, the best-fit $PE_{delayed}$ count is consistent with the *Fast Simulation* result, which demonstrates that the *Fast Simulation* is applicable not only for the special event hypothesis validated in Section 7.2.1, but also for a realistic event composition.
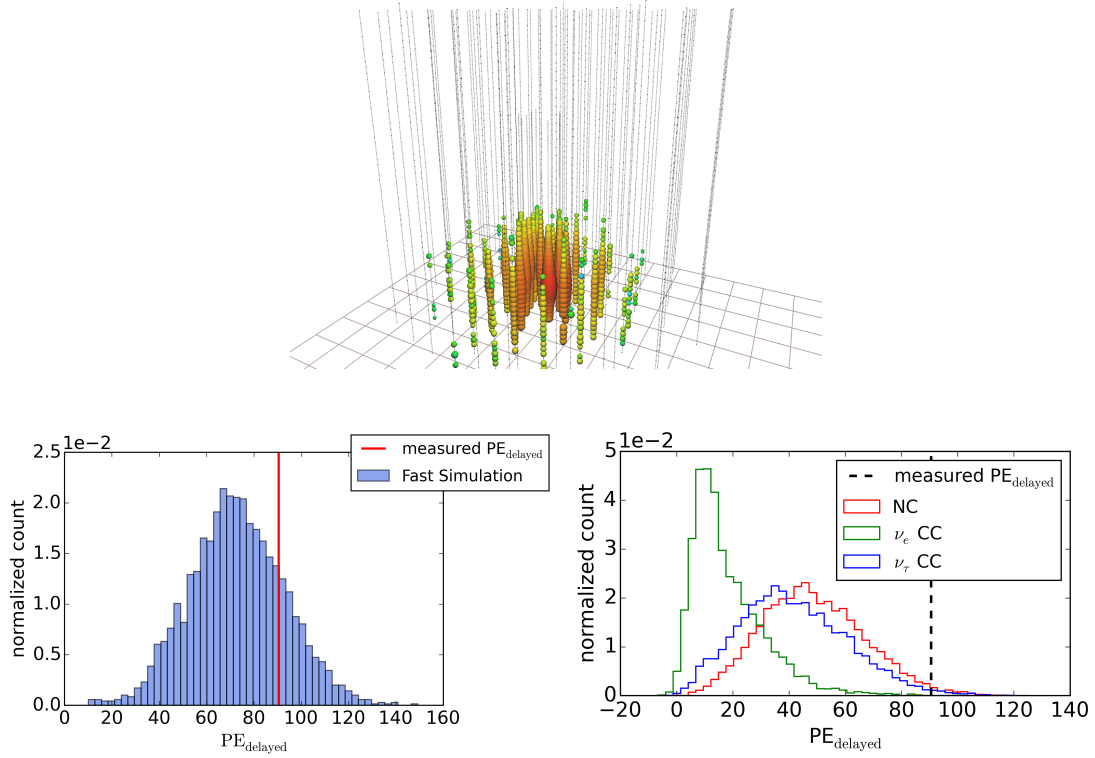
Figure 7.20: Simulated ν_τ CC shower, 168 TeV reconstructed energy. An event view of the simulated cascade is shown at the top. The lower left and lower right distributions are produced by the *Fast Simulation* for a known and unknown shower content, respectively. The red line in the lower left plot and the black dashed line in the lower right plot mark the event's best-fit $PE_{delayed}$ value.

7.6.2 Expected event rates for the analysis dataset

The high-statistics HESE simulation dataset was furthermore used to determine the expected number of events and interaction type composition in the analysis dataset. The analysis dataset will contain all shower-like HESE events collected in 22 months of livetime with a charge greater than 6000 PE and more than 60 TeV reconstructed energy.

The expected event rates from atmospheric neutrino fluxes are listed in Table 7.2. No tau flavor is predicted in the atmospheric component, assuming that the contribution from the prompt process is small. With less than 0.8 expected events, the atmospheric muon background can be neglected.

76

flux	NC	CC, cascades only	CC, total
conv. atmo. ν_e	0.012 ± 0.001	0.263 ± 0.001	0.263 ± 0.001
conv. atmo. ν_μ	0.378 ± 0.015	-	1.777 ± 0.035
conv. atmo. ν_τ	-	-	-
prompt atmo. ν_e	< 0.042	< 0.613	< 0.613
prompt atmo. ν_μ	< 0.043	-	< 0.216
prompt atmo. ν_τ	-	-	-

Table 7.2: Event rates expected from atmospheric fluxes in 22 months of HESE data with more than 6000 PE charge deposition and more than 60 TeV reconstructed energy.

The astrophysical contribution was calculated for four different assumption for the spectral index γ in an unbroken power law[10] and assuming a 1 : 1 : 1 flavor mixture at Earth. The astrophysical event rates are summarized in Table 7.3.

	flux	NC	CC, cascades only	CC, total
	astro. ν_e	0.663 ± 0.004	4.390 ± 0.008	4.390 ± 0.008
$\gamma = 2.0$	astro. ν_μ	0.679 ± 0.008	-	2.533 ± 0.016
	astro. ν_τ	0.690 ± 0.008	3.005 ± 0.017	3.426 ± 0.019
	astro. ν_e	0.621 ± 0.003	5.488 ± 0.010	5.488 ± 0.010
$\gamma = 2.3$	astro. ν_μ	0.633 ± 0.008	-	2.525 ± 0.017
	astro. ν_τ	0.632 ± 0.008	3.298 ± 0.010	3.691 ± 0.021
	astro. ν_e	0.598 ± 0.003	6.603 ± 0.013	6.603 ± 0.013
$\gamma = 2.58$	astro. ν_μ	0.609 ± 0.009	-	2.528 ± 0.018
	astro. ν_τ	0.598 ± 0.009	3.585 ± 0.023	3.960 ± 0.024
	astro. ν_e	0.449 ± 0.003	6.224 ± 0.014	6.224 ± 0.014
$\gamma = 2.92$	astro. ν_μ	0.457 ± 0.008	-	1.961 ± 0.017
	astro. ν_τ	0.442 ± 0.007	3.028 ± 0.022	3.307 ± 0.023

Table 7.3: Event rates expected from astrophysical fluxes in 22 months of HESE data with more than 6000 PE charge deposition and more than 60 TeV reconstructed energy.

---

[10]The investigated spectral indices correspond to the naive expectation derived in the *first order diffusive shock acceleration* model and the best-fit values from different IceCube analyses (see Chapter 4).

Translating the flux event counts to shower type expectations finally yields the event counts summarized in Table 7.4.

	NC	ν_τ CC	ν_e CC	fraction NC
$\gamma = 2.0$	$2.422^{+0.062}_{-0.019}$	3.005 ± 0.017	$4.653^{+0.613}_{-0.008}$	24 %
$\gamma = 2.3$	$2.276^{+0.062}_{-0.019}$	3.298 ± 0.012	$5.751^{+0.650}_{-0.010}$	20 %
$\gamma = 2.58$	$2.195^{+0.063}_{-0.023}$	3.585 ± 0.020	$6.866^{+0.613}_{-0.013}$	17 %
$\gamma = 2.92$	$1.738^{+0.063}_{-0.019}$	3.028 ± 0.022	$6.487^{+0.613}_{-0.014}$	15 %

Table 7.4: Expected event rates and interaction type composition of the analysis dataset.

Typically, only $\mathcal{O}(10\,\%)$ of the neutrino energy is transferred to the target nucleus in a neutrino NC interaction. The neutrino energy thus needs to be around 10 times larger than the energy of the observed hadronic cascade. A larger astrophysical index consequently reduces the NC fraction in the data sample as it leads to a softer astrophysical energy spectrum.

7.7 Statistical methods to characterize cascade types

In this section, the statistical prerequisites for the evaluation of the individual cascade type as well as the interaction type composition of the ensemble of showers will be given. The presentation style of the ensemble result will be demonstrated by means of the HESE simulation dataset introduced in the previous section.

7.7.1 Individual event anaylsis

For every event in the analysis dataset, information allowing for a statement on the compatibility of the observed shower with a certain interaction type will be presented. The presentation style is demonstrated here by means of a 140 TeV simulated NC shower. Figure 7.21 shows the $PE_{delayed}$ distributions from the *Fast Simulation* with the event's best-fit $PE_{delayed}$ value marked.
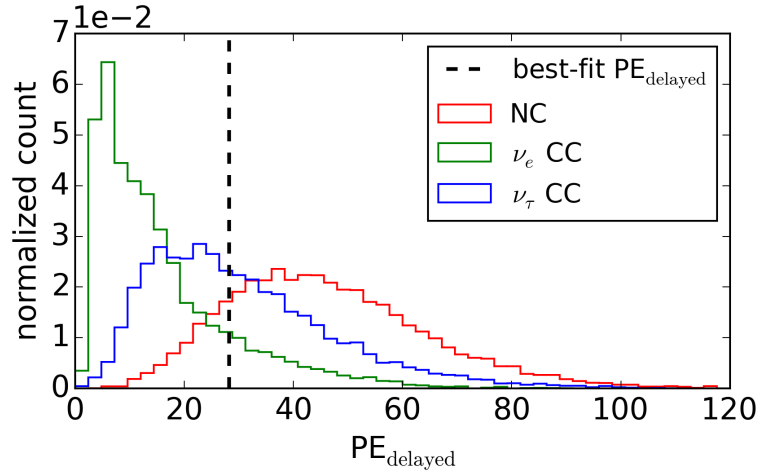
Figure 7.21: $PE_{delayed}$ distributions. The dashed line marks the best-fit $PE_{delayed}$ for the simulated NC event.

In addition, the value of the inverse cumulative distribution function (iCDF)[11] for the three distributions at the best-fit $PE_{delayed}$ will be given. For the example NC shower discussed above and the probability density distributions in Figure 7.21 one would obtain:

	iCDF($PE_{delayed,fit}$)
ν_e CC	0.17
ν_τ CC	0.52
NC	0.87

To demonstrate the impact of the systematic uncertainties, cumulative distribution functions (CDFs) corresponding to the $PE_{delayed}$ distributions will be shown together with bands illustrating the $\pm 56\,\%$ uncertainty range.

---

[11]Cumulated from the right.

(a) ν_e CC   (b) ν_τ CC   (c) NC
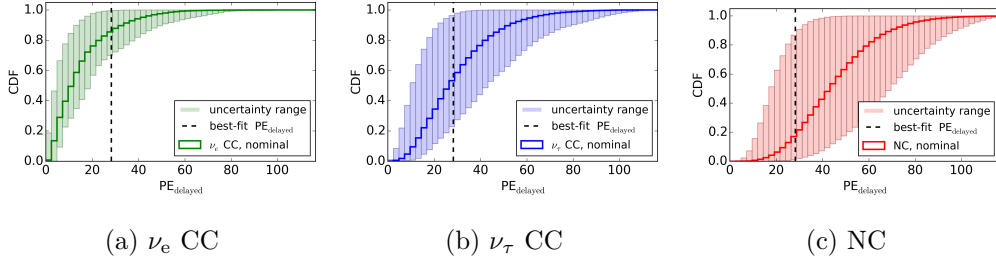
Figure 7.22: CDFs corresponding to the distributions in Figure 7.21. The bands show the $\pm 56\,\%$ systematic uncertainty range.

7.7.2   Statistical interpretation of the ensemble: Frequentist approach

In addition to the individual event information, a likelihood fit is used to determine the interaction type composition of the *ensemble* of events.

Let α and β be the relative NC and ν_e CC contribution, respectively. Then, the ν_τ CC contribution will be given by $(1 - \alpha - \beta)$. When d denotes an event's best-fit $\mathrm{PE_{delayed}}$ value, the per-event likelihood function is given by:

$$\mathscr{L}_i\,(d\,|\,\alpha, \beta) \;=\; \alpha \cdot P_{NC}(d) \;+\; \beta \cdot P_{\nu_e\,CC}(d) \;+\; (1 - \alpha - \beta) \cdot P_{\nu_\tau\,CC}(d) \quad , \tag{7.8}$$

where $P_i(d)$ is the value of the respective probability density function (histograms in Figure 7.21) for $\mathrm{PE_{delayed}} = d$ (black dashed line in Figure 7.21).

To include the 56 % uncertainty on n_m^d in Equation 7.2 as a nuisance parameter[12] s, the dependency of $P_i(d)$ on s needs to be incorporated. Therefore, the n_m^d counts contributing to the $\mathrm{PE_{delayed}}$ distributions are scaled according to

$$n_m^{d\,\prime} = n_m^d + s \cdot n_m^d \tag{7.9}$$

and subsequently smeared to account for the analysis' relative resolution (Figure 7.12). $P_i(d)$ can then again be read off from the modified distribution. Repeating the procedure for different values of s yields $P_i(d, s)$ distributions as displayed in Figure 7.23 for a showcase ν_τ CC event. To provide smooth input functions to the likelihood fit, an exponentially modified Gaussian[13] is fitted to the distributions. The likelihood function

---

[12]A commonly used method to include systematic uncertainties in a likelihood fit is the implementation of additional, free parameters called *nuisance parameters*. The fit is then allowed to find the parameter values describing the data best. Prior knowledge on the uncertainty of the nuisance parameters can be incorporated e.g. through Gaussian penalty factors, imposing the uncertainty range on the likelihood landscape.

[13]The probability density function of the exponentially modified Gaussian distribution is given by: $f(x; \mu, \sigma, \lambda) = \frac{\lambda}{2} \cdot \exp\left(\frac{\lambda}{2} \cdot (2\mu + \lambda\sigma^2 - 2x)\right) \cdot \mathrm{erfc}\left(\frac{\mu + \lambda\sigma^2 - x}{\sqrt{2\sigma^2}}\right)$.

for the ensemble of events, including s as a nuisance parameter, finally reads:

$$\mathscr{L}(d \,|\, \alpha, \beta, s) = \prod_i \mathscr{L}_i(d \,|\, \alpha, \beta, s) \times e^{-\frac{1}{2}\left(\frac{s-0}{\sigma_s}\right)^2} \quad , \tag{7.10}$$

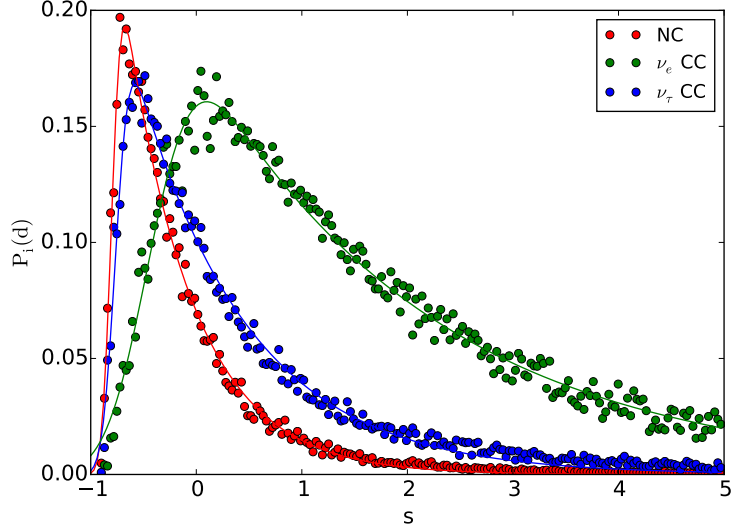where σ_s is the $56\,\%$ uncertainty on the echo strength discussed in Section 7.5.

Figure 7.23: $P_i(d, s)$ distributions for a sample ν_τ CC event. The solid lines show the corresponding best-fit exponentially modified Gaussian.

The best-fit result for the cascade sample will be presented in a triangle plot familiar from IceCube's global flavor fit presented in Chapter 4. In this analysis, however, the axes are given by the different shower types instead of neutrino flavors. The triangle plot for the 13 event in the HESE simulation dataset is displayed in Figure 7.24. Therein, a likelihood ratio test (see Section 7.3) is used to determine the confidence regions for the interaction type composition. The test statistic at a test point with hypothesis H_0 is the ratio of the corresponding likelihood value over the likelihood value at the best-fit hypothesis H_i:

$$TS \equiv -2\Delta \log(\text{Likelihood}) = -2 \cdot \ln\left(\frac{\mathscr{L}_{H_0}}{\mathscr{L}_{H_i}}\right) \quad . \tag{7.11}$$

At every test point in the triangle, the nuisance parameter s is re-optimized. Given the computational requirements of the analysis, producing radomized test statistic distributions for the determination of the confidence regions via the Toy Monte Carlo method is not feasible. Instead, *Wilk's theorem* is applied which predicts that, under suitable regularity conditions and for large sample sizes, the test statistic distribution of null hypothesis realizations follows a χ^2 distribution with degrees of freedom equal to the
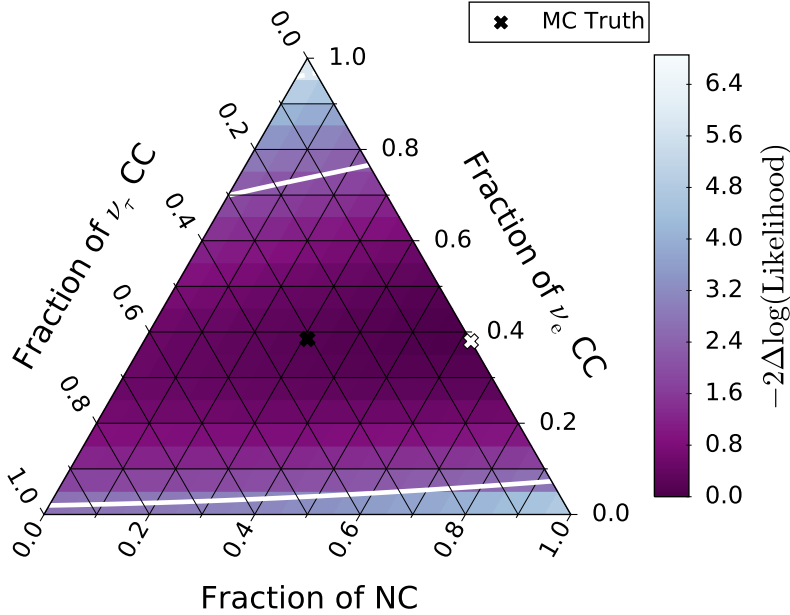
Figure 7.24: Two-dimensional likelihood scan around the best-fit point (white cross, [NC, ν_e CC, s] = [0.62, 0.38, 0.22]) of the representative simulation dataset. The black cross marks the true Monte Carlo input (MC Truth). The solid and dashed white lines refer to the 68 % and 95 % confidence intervals, respectively.

difference of free parameters in $\mathscr{L}_{H_i}$ and $\mathscr{L}_{H_0}$ [96]. Since the NC and ν_e CC contributions are fixed at every test point, $\mathscr{L}_{H_0}$ has only s as a free parameter left and a χ^2 distribution with $3 - 1 = 2$ degrees of freedom is therefore appropriate. The xth percentile of the χ^2 distribution then marks the x % confidence interval.

Given the computational limitations, Wilk's theorem lets one conveniently calculate confidence regions, although, with around eleven expected events[14], the application requirement of a large sample size is not met. Whether the test statistic is nevertheless χ^2 distributed was tested by means of 10,000 pseudo experiments. The total number of events in each Toy Monte Carlo run was randomized following a Poisson distribution with mean value eleven. The randomized number of events was subsequently drawn from a pool of events, consisting of showers at around 100 TeV in the expected interaction type composition. The resulting test statistic distribution in Figure 7.25 shows that a χ^2 distribution with two degrees of freedom predicts rather conservative confidence intervals. Although not reflecting the energy and vertex distribution of the final dataset, the Toy Monte Carlo setup is considered a representative scenario and Wilk's theorem is applied in the data analysis.

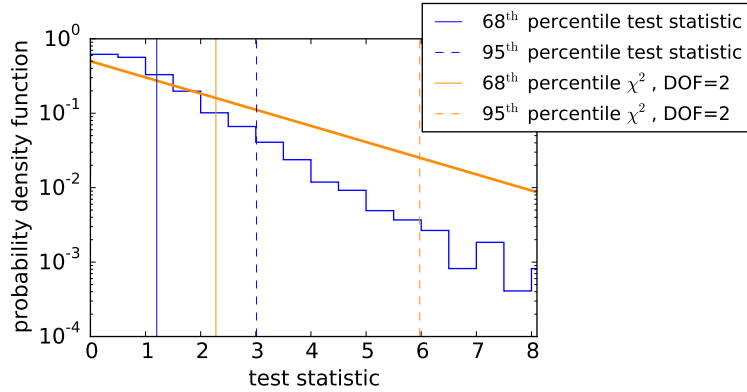[14]13 events were actually found in the analysis dataset.

Figure 7.25: A χ^2 distribution with two degrees of freedom (orange curve) predicts conservative confidence intervals given the test statistic distribution of the Toy Monte Carlo experiments (blue histogram).

7.7.3 Statistical interpretation of the ensemble: Markov Chain Monte Carlo approach

Figure 7.25 shows that the confidence regions calculated using Wilk's theorem run the risk of being overly conservative in this analysis. In addition, the small data sample does not fulfill the theorem's prerequisite of a large sample size. To add statistical robustness to the results calculated in the frequentist approach, quantities analogous to the best-fit point and confidence regions are in addition determined via the *Markov Chain Monte Carlo* (MCMC) method.

The MCMC approach for optimization problems in computational physics is growing in popularity, especially for models with a large number of free parameters. The aim of a MCMC is to sample from the *posterior probability distribution function* $p(\Theta|x)$ of the parameter Θ given the data x. Using Bayes' theorem [97], $p(\Theta|x)$ can be linked to the likelihood function $p(x|\Theta)$ via the relation

$$p(\Theta|x) \;=\; \frac{p(x|\Theta)\, p(\Theta)}{p(x)} \qquad , \tag{7.12}$$

where $p(x)$ is the probability to measure x and $p(\Theta)$ is the prior for the probability distribution function of Θ. The parameter Θ can of course be multi-dimensional. A method, such as MCMC, able to sample from $p(\Theta|x)$, can provide an approximation for the posterior probability distribution function, which can then be used to determine the most probable parameter Θ given the data x as well as its credible regions.

To obtain a sample approximation for $p(\Theta|x)$, so-called *walkers* explore the parameter space in a random walk. In the fashion of a Markov chain, a walker's next step only depends on its current position. In the MCMC routine used in this work (*emcee* [98]), walkers are moved sequentially and a suggestion for the walker's new position is calculated taking the positions of the remaining walkers into account. Whether a proposed move from Θ_t to Θ_{t+1} is accepted is determined using the *Metropolis-Hastings rule*[15], in which the acceptance probability is proportional to $p(\Theta_{t+1}|x)/p(\Theta_t|x)$. Consequently, walkers will step to points yielding a higher value in the posterior probability distribution function more often. One can thus deduce the most probable Θ and its credible regions from the density distribution of the parameter points the walkers moved to.

The likelihood function in Equation 7.12 is the same as the one used in the frequentist approach, except for an individual fit parameter γ for the ν_τ CC fraction, implemented to test whether the implicit definition in Equation 7.10 causes a bias. For the prior $p(\Theta)$, a uniform distribution, or *uninformative* prior, is used. $p(x)$ is independent of Θ and does not need to be computed explicitly [98].

Figure 7.26 shows the one-dimensional histograms of the sampled points for each model parameter (NC, ν_e CC, ν_τ CC, s) for the 13 showers in the HESE simulation dataset, as well as their two-dimensional projections[16]. The most probable value for the nuisance parameter s is 0.26, which is in agreement with the likelihood best-fit value of 0.22.

In analogy to the two-dimensional likelihood scan in the frequentist approach, the sample density is displayed in a triangle plot in Figure 7.27. Here, the $y\,\%$ credible interval is given by $[\Theta_1, \Theta_2]$, if the interval fulfills the condition

$$\int_{\Theta_1}^{\Theta_2} p(\Theta|x)\,d\Theta = y\,\% \qquad . \tag{7.13}$$

Only the largest values of $p(\Theta|x)$ are included in the interval. This ensures that the interval covers the highest density point. Although the interpretation of Bayesian credible intervals and frequentist confidence intervals is rather different, they are often numerically identical [102]. Assuming this is also the case in this work, the credible regions in Figure 7.27 demonstrate that the confidence regions calculated by means of Wilk's theorem (see Figure 7.24) are too conservative. The point of the highest sample density ([NC, ν_eCC, ν_τCC] = [0.55, 0.45, 0.0]) is consistent with the likelihood best-fit value ([NC, ν_eCC, ν_τCC] = [0.62, 0.38, 0.0]).

---

[15]The algorithm is named after N. Metropolis, author of the first paper covering the algorithm [99], and W. Hastings, who later generalized the routine [100].

[16]Figure 7.26 is an example of a so-called *corner plot* - a scheme conveniently illustrating the sample density of an MCMC. In this work, the *corner* module [101] is used to create corner plots.
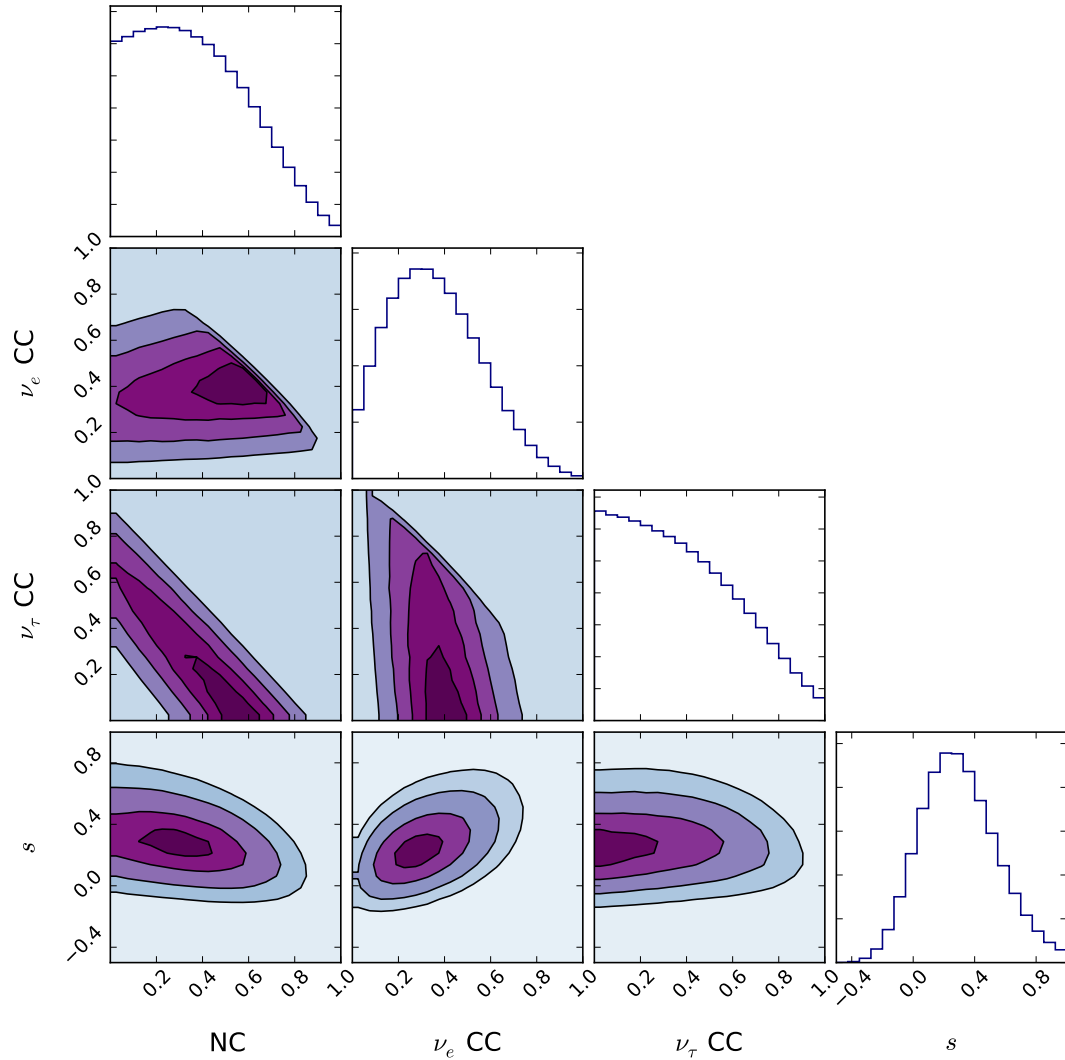
Figure 7.26: Corner plot of the MCMC result for the HESE simulation dataset. The filled areas in the two-dimensional density plots refer to the $1\,\sigma$, $2\,\sigma$, $3\,\sigma$ and $4\,\sigma$ region. The boundary conditions $\alpha, \beta, \gamma \in [0,1]$ and $\alpha+\beta+\gamma = 1$ reject the upper right parameter space in the two-dimensional projections of the relative interaction type contributions. The allowed range for α, β and γ has been slightly extended to $[-0.1,1.1]$ in the MCMC to avoid edge effects.
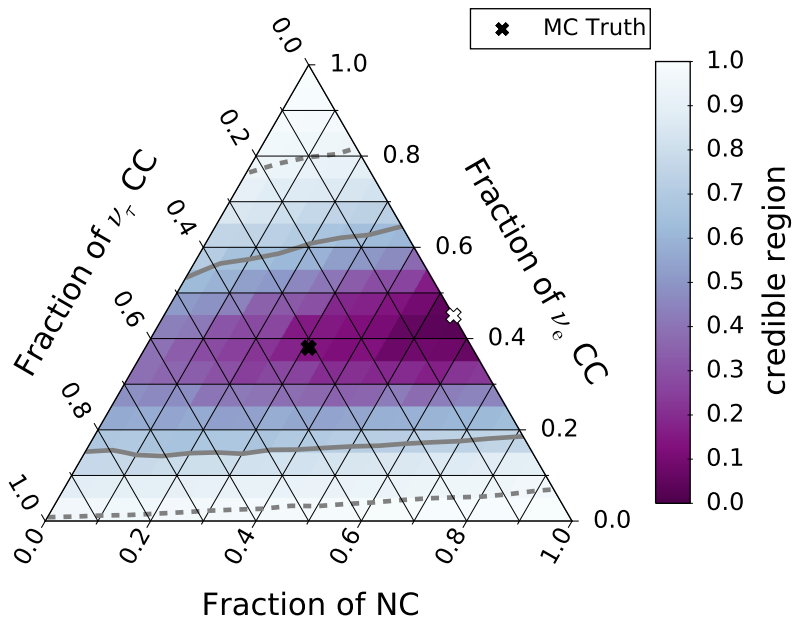
Figure 7.27: The triangle plot reflects the two-dimensional sample density of the MCMC for the 13 simulated HESE events. The solid and dashed gray lines mark the 68 % and 95 % credible regions, respectively. The highest density point, which is the analog to the best-fit point of a likelihood fit, is shown as a white cross.

7.8 Test for consistency with the standard model scenario and beyond the Standard Model physics

The ensemble fit result will be tested with regard to its consistency with the *standard* hypothesis, which sees IceCube's high energy events originating from atmospheric and astrophysical neutrino fluxes and assumes that astrophysical neutrinos are produced with a flavor ratio of $1 : 2 : 0$ in $\nu_e : \nu_\mu : \nu_\tau$ and oscillate according to the Standard Model on their journey to the Earth. Here, this scenario is referred to as the *standard model hypothesis*. In addition, the ensemble will be used to probe the assumption that the detected high energy cascades are produced in the scattering of dark matter particles, as suggested by two exotic theories.

7.8.1 Probing the standard model scenario

The favored explanation for the origin of HESE events is an astrophysical neutrino flux (see Table 7.3) in addition to the atmospheric component (see Table 7.2). As the astrophysical flux hypothesis, an unbroken power law with a $1 : 1 : 1$ flavor composition (after long baseline oscillation) at the detector is assumed. For the power law spectral index, the best-fit results for the three, four and six year HESE data and $\gamma = 2$, the naive expectation for a production via diffusive shock acceleration, will be tested (see Section 4.3). Figure 7.28 again shows the shower type triangle for the 13 simulated HESE events, but with the average interaction type composition for the standard model assumptions additionally marked. In the frequentist approach, one can calculate a p-value for each of the points, which states the consistency of the standard model null hypothesis with the best-fit result. For the 13 representative simulation events, one obtains:

assumed astrophysical index	p-value
2	0.82
2.30	0.73
2.58	0.67
2.92	0.59

Since the simulation set was produced under the standard model hypothesis with $\gamma = 2.3$, the ensemble's best-fit results are consistent with the standard model assumption.

A Bayesian measure comparable to the frequentist p-value is the edge of the credible region a point lies on. In contrast to a p-value of $x\,\%$, which rejects the null hypothesis at a $(100-x)\,\%$ confidence level, a point's $x\,\%$ Bayesian credible region lets one conclude that, given the measured data and model hypothesis, the probability of observing this point is $(100-x)\,\%$. For the HESE simulation dataset, the observation probabilities of the standard model hypotheses are:
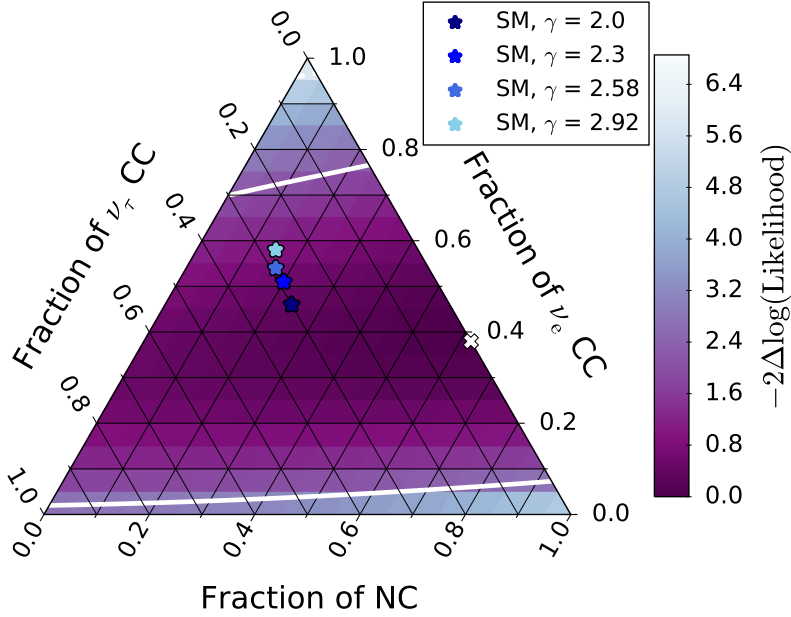
Figure 7.28: In addition to the two-dimensional likelihood scan in Figure 7.24, blue markers show the average shower type composition for the standard model (SM) hypothesis with different assumptions for the astrophysical index. The atmospheric flux assumption is the same in all scenarios (see Table 7.2).

assumed astrophysical index	observation probability
2	0.68
2.30	0.52
2.58	0.42
2.92	0.33

Again, the input value of $\gamma = 2.3$ is consistent with the MCMC result; however, the small sample of 13 cascades does not allow for a distinction between the spectral index hypotheses.

7.8.2 Probing boosted dark matter scenarios

The nature of the universe's dark matter content is a major puzzle of modern physics. Assuming that dark matter is made of one or multiple yet unknown particle species, rather small-sized but very sensitive experiments usually look for a recoil signature from dark matter scatterings on nuclei (direct detection), while the larger neutrino observatories search their data for an excess of detected neutrinos produced in dark matter decay or annihilation (indirect detection). There are, however, alternative approaches proposing to use neutrino observatories as huge direct detection experiments. To pro-

duce a detectable signature in IceCube, the interacting dark matter particles need to be extremely energetic. Assuming a very heavy dark matter species, which decays or annihilates into a second much lighter dark matter component, would yield such relativistic particles as the final state dark matter particles would be *boosted* due to the mass difference. The production of boosted dark matter particles via the decay or annihilation of a heavy dark matter species is sketched in Figure 7.29. While the boosted dark matter component can interact with Standard Model particles, the heavy species would have no interactions with the Standard Model.

Several authors interpret IceCube's HESE showers as being solely or partially due to deep inelastic scatterings of boosted dark matter particles with nuclei in the ice [103–105]. The nuclear debris would yield an event signature indistinguishable from the shower from a neutrino neutral current interaction (see Figure 7.29 (a)).

Other models suggest a more *leptophilic* boosted dark matter interaction, in which the hypothetic particles elastically scatter on atomic electrons [106, 107]. The high energy electron produced in the interaction would initiate an electromagnetic shower in analogy to the electron (positron) generated in electron neutrino (antineutrino) charge current interactions. This scattering process is illustrated in Figure 7.29 (b).
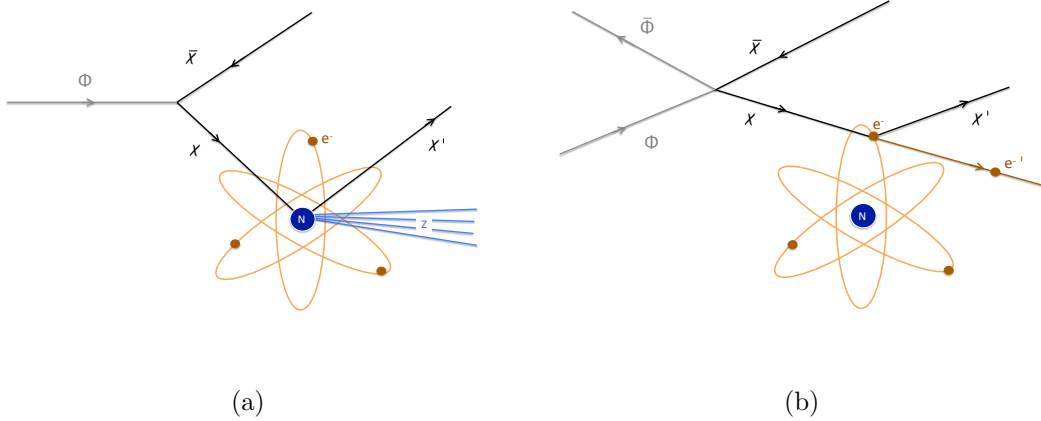

(a)   (b)

Figure 7.29: The left-hand sketch illustrates the DIS of a boosted dark matter particle (χ), generated in the decay of a much heavier dark matter particle (Φ), on a nucleus (N). In analogy to the DIS of neutrinos, the nuclear debris (Z) produced in the interaction induce a hadronic shower. On the right-hand side, the elastic scattering of a boosted dark matter particle (χ), produced via annihilation of the heavy dark matter species (Φ), on an atomic electron is shown. The released high energy electron leads to an electromagnetic cascade in analogy to the e$^\pm$ generated in $\overset{(-)}{\nu}_e$ CC interactions.

Boosted dark matter models have at least six free parameters[17] and would lead to a mono-energetic excess of one shower type in the data. Due to the complexity of the models and the shower energies in the analysis dataset, there is so far no concrete model proposed that can be tested in this work. Nevertheless, an excess of solely electromagnetic or hadronic showers would be a clear hint at beyond the Standard Model physics and the presented scenarios offer a compelling incentive to explicitly probe the 100% NC and 100% ν_e CC assumption.

The rejection level for the 100% NC and 100% ν_e CC assumption can be read off from the triangle in Figure 7.28. For the 13 HESE simulation events, the 100% NC and 100% ν_e CC hypothesis is rejected at 92% and 97% confidence level, respectively. In addition, one-dimensional likelihood scans allow the determination of the maximum NC and ν_e CC fraction consistent with the data. In contrast to the confidence level calculation in the triangle plot, only the NC or ν_e CC contribution is fixed in the likelihood at every test point and the remaining parameters are re-optimized. The difference in the degrees of freedom in the alternative hypothesis (best-fit point) and the null hypothesis (test point) is thus $3 - 2 = 1$. Following Wilk's theorem, a χ^2 distribution with one degree of freedom is used for the confidence level calculation. Figure 7.30 shows the one-dimensional likelihood scans for the NC and ν_e CC fraction in the HESE simulation dataset. At 90% confidence level, the maximum allowed relative NC and ν_e CC contribution is 92% and 80%, respectively.
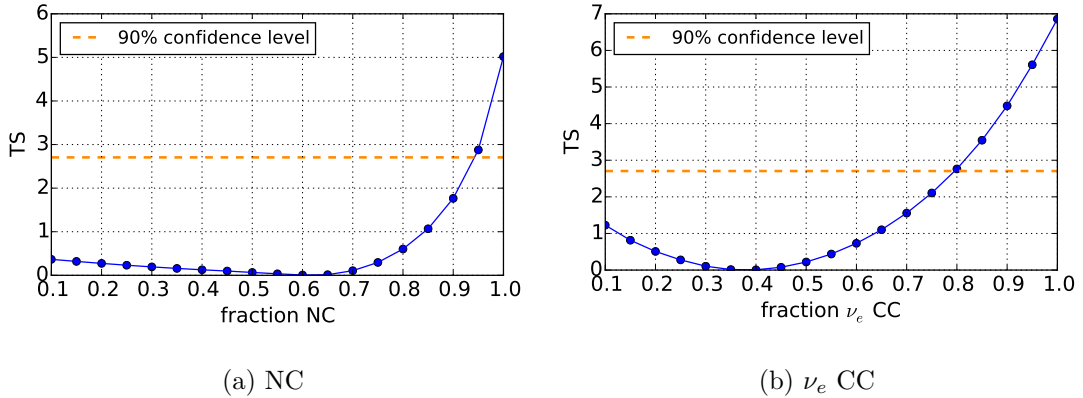

(a) NC

(b) ν_e CC

Figure 7.30: One-dimensonal likelihood scans for the HESE simulation dataset. Applying Wilk's theorem, a χ^2 distributions with one degree of freedom determines the 90% confidence level.

---

[17]Free model parameters are - at least - the mass of the heavy and light dark matter particle, the annihilation or decay cross section, the mass of the particle that mediates the Standard Model-dark matter interaction (e.g. the *dark photon*) and the coupling strengths between the mediator particle and the boosted dark matter and Standard Model particle.

The large confidence regions in the triangle plot reflect the large systematic uncertainties and the so far small sample size. However, even with low statistics and large uncertainties, the extreme hypothesis of 100 % NC or ν_e CC contribution, which would hint at physics beyond the Standard Model, can already be excluded beyond the 90 % confidence level in the HESE simulation set.

The MCMC disfavors the extreme scenarios even stronger. The 100 % NC as well as the 100 % ν_e CC point lies outside the 99 % credible region, meaning that the probability to measure these parameter points is less than 1 % for the HESE simulation dataset.

8

Analysis of the HESE HitSpool dataset

In this chapter, the analysis method developed in Chapter 7 is applied to the analysis dataset. Results are presented on an event-by-event basis and the ensemble of showers is analyzed. The determined interaction type composition of the data sample is studied with regard to its consistency with the standard model hypothesis and furthermore interpreted within two boosted dark matter concepts. At first, a delayed excessed observed for HESE alerts with less than 6000 PE is studied in more detail.

8.1 The HESE HitSpool dataset with 1500-6000 PE

In addition to the high energy HESE selection with a threshold of 6000 PE, HitSpool data are collected for HESE alerts in the 1500-6000 PE range. This low-PE dataset is a very unclean sample in the sense that it is composed of around 97 % atmospheric muons (see Figure 4.2). The muon contamination can, however, be exploited to study muonic stochastic energy losses, once the uncertainties in the neutron echo analysis become smaller. With smaller uncertainties, the sample would furthermore present an excellent calibration tool for the neutron echo itself.

8.1.1 Delayed excess in the stacked PE spectrum

In order to guarantee the analysis' blindness[1], only HESE events below the 6000 PE threshold were studied in advance. In this low-PE data, an excess in the stacked launch spectrum[2] of around 1200 events, collected from March 2016 to April 2017, at around $100\,\mu$s is nevertheless visible.

Whether the observed signal is consistent with the neutron echo hypothesis was investigated by means of a binned Poisson likelihood fit. When d_i and f_i denotes the experimental and expected data in bin i, respectively, the likelihood function is given by:

$$\mathscr{L}(d\,|\,f) = \prod_i^{\#bins} \frac{f_i^{d_i}}{d_i}\,e^{-f_i} \qquad . \tag{8.1}$$

---

[1]To avoid a possible bias, analysis methods must not be developed or tested on their dedicated datasets. Every analysis method is inspected in an IceCube internal review process before it can run on real data.

[2]Due to problems in the reconstruction of delayed pulses in the long HitSpool readout window, the analysis had to be based on DOM launches instead of pulses. See Section 8.2 for details.

For each event, the time dependent deadtime correction function was calculated by the routine introduced in Section 7.2 and combined to an average deadtime correction for the ensemble. As the neutron echo is described by an exponential shape, the expectation values f_i under the echo hypothesis are given by the deadtime-corrected exponential distribution with the amplitude and decay time τ as free parameters.

Studying the consistency of the observed excess with the neutron echo assumption, two questions arise: Does the observed shape match a neutron capture induced signal? Is the strength in agreement with the neutron echo hypothesis? The first question is answered by the best-fit value for the distribution's decay time τ. The likelihood fit finds a low value of $\tau = 166^{+23}_{-20}\,\mu s$ to describe the data best. A decay time of around $217\,\mu s$, expected from Geant4 shower simulations (see Chapter 6), is, however, still compatible with the data at the $2\,\sigma$ a level. It is difficult to imagine how the capture time - which is well measured in water ($(203.7\pm2.8)\,\mu s$ [108]) - can be lower in high purity ice.

For the expected signal strength, the following considerations and assumptions go into the estimation:

- The average reconstructed energy in the dataset is around $10\,\mathrm{TeV}$.

- The low-PE HESE sample consists of $97\,\%$ atmospheric muons. All atmospheric muons that passed the HESE veto experienced catastrophic energy losses.

- At $10\,\mathrm{TeV}$ visible energy, $20\,\%$ of the muonic stochastic losses occur via photonuclear reactions (see Figure 2.5), which initiate a hadronic cascade. Thus, around $20\,\%$ of the events are expected to be NC-like.

- Figure 8.1 shows that $0.18\,\%$ of the prompt light is observable in the echo of $\mathcal{O}(100\,\mathrm{TeV})$ NC showers. The relative echo strength in $10\,\mathrm{TeV}$ cascades is expected to be $4\,\%$ stronger due to the larger hadronic shower content at lower shower energies (see Figure 2.3).

The prompt light of the stacked PE spectrum contains roughly $2.3{\cdot}10^6$ PE. Consequently,

$$2.3 \cdot 10^6\,\mathrm{PE} \times 20\% \times 0.18\% \times 1.04 \approx 860\,\mathrm{PE} \tag{8.2}$$

are expected in the delayed signal. The likelihood fit finds around $1250\,\mathrm{PE}$ contained in the excess when assuming that it can be attributed to the neutron echo.

In summary, the decay time of the delayed excess in the low-PE HESE sample is consistent with the expectation of $217\,\mu s$ at the $2\,\sigma$ level, while the echo strength is around 1.5 times larger than estimated by the back-of-the-envelope calculation discussed above. It must, however, be emphasized that the low-PE sample was only intended as a pre-unblinding test ground and that the uncertainties involved are not well under control yet. For example, a factor two uncertainty is already associated with the echo strength itself and the muon photonuclear rate is only known with an approximately $20\,\%$ error

[109]. The best-fit spectrum in Figure 8.2 furthermore shows that the averaged deadtime function is not well suited for the description of the first 150 μs. A dedicated IceCube simulation set, ideally including Geant4 particle tracking for the muon's reaction products, would be needed to study the type of muon events passing the HESE veto, the muon energy losses and their neutron echo and the precise deadtime effects. Producing such a simulation set is, however, extremely time consuming and CPU intense[3] and thus beyond the scope of this work. This thesis may state that there is an undeniable excess in the delayed time spectrum that has never been measured in IceCube before. The observed excess is compatible with the neutron echo hypothesis and large efforts are needed to advance the echo analysis of the low-PE sample. This observation was taken as a strong motivation to continue and advance the neutron echo analysis.
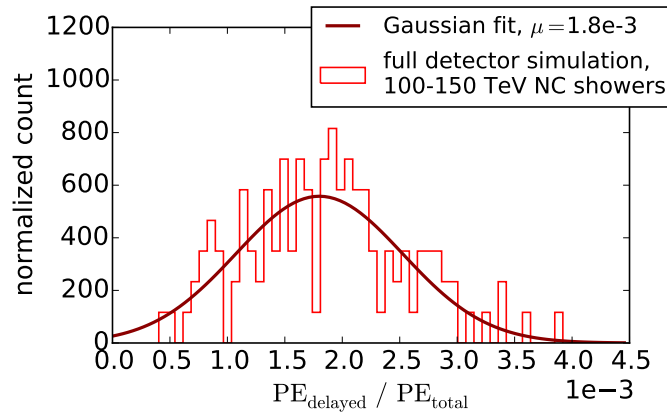
Figure 8.1: The delayed signal relative to the prompt light is shown for 100-150 TeV NC showers simulated using the full detector simulation.

---

[3]Assuming the same computational requirements as for the performed full detector cascade simulation, around 500,000 CPU hours are needed to track the reaction products of $\mathcal{O}(1000)$ muon events. This estimate does not yet include the computation time required for the simulation of the muon interactions and the preselection of events with a catastrophic energy loss in the detector.
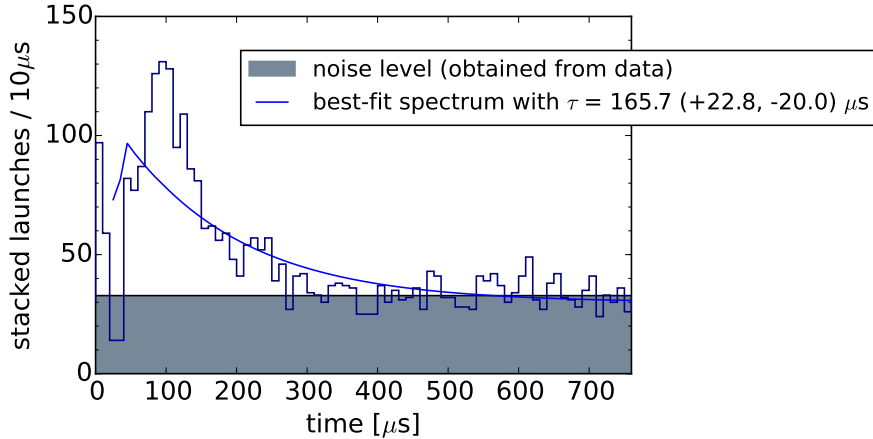
Figure 8.2: The stacked launch spectrum of around 1200 low-PE HESE events together with the best-fit spectrum under the neutron echo hypothesis. The noise expectation is directly extracted from the 0.5 s of data recorded before the actual event. Note that the prompt peak (first bins) is significantly higher than the echo distribution when considering pulses instead of launches.

8.1.2 Delayed light from low energy muons

No delayed light signal is expected from muons with energies below around 1 TeV (see Section 2.3.2). A study of the light emission subsequent to the passage of a low energy muon thus provides a good test of whether a yet unknown source, correlated with light deposition in the detector, causes the observed delayed excess discussed in the previous section.

Low energy muons, contained in the 0.5 s of HitSpool data before the start of a HESE event, were investigated. Hits are assigned to a muon track when they were tagged as muon hits in the HitSpool processing chain (see Section 5.2) and additionally belong to a hit cluster identified by the SMT8 trigger[4]. Figure 8.3 shows the time delay of hits in a DOM subsequent to a hit from an SMT8 muon. As expected, no delayed light emission is observable. The excess in the first bin originates from scattered light. The data are slightly below the 560 Hz noise level at low delay times due to deadtime effects. The delay time distribution for pure noise hits in Figure 8.3 reveals the correlated noise contribution discussed in Section 3.2.1.

---

[4]The trigger vetoes noise by requiring at least eight HLC hits within 5 μs and is therefore named *Simple Multiplicity Trigger 8* or SMT8.
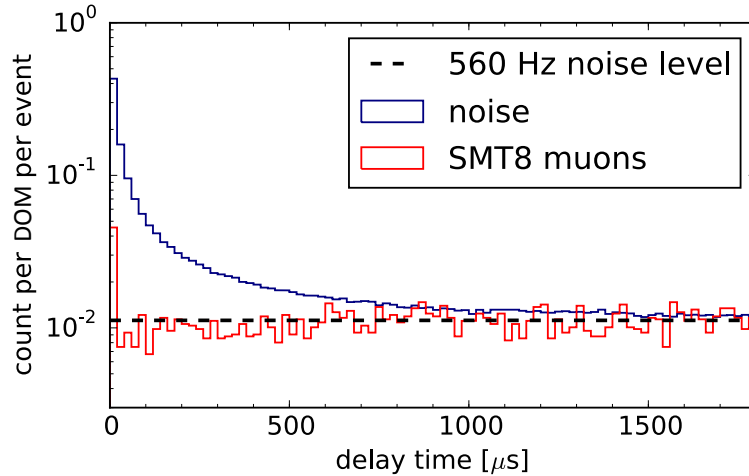
Figure 8.3: The delay time distribution of hits following a low energy muon that fulfilled the SMT8 criterion is shown in red. Scattered light from the muon event itself causes an excess in the first bin of the muon delay time spectrum. The delay times of pure noise hits (purple histogram) illustrate the correlated noise contribution.

8.2 The dataset used for the final analysis

The analysis dataset consists of HitSpool data for all shower-like HESE events collected between February 25, 2016 and December 31, 2017 (676 days) with a charge greater than 6000 PE and more than 60 TeV reconstructed energy. Before discussing the analysis results, a problem in the extraction of delayed pulses, which was only discovered after analyzing the final dataset, is discussed.

8.2.1 Faulty delayed pulse reconstruction

The delayed PE spectrum of the HESE HitSpool dataset revealed several ATWD waveforms from which a multitude of pulses was reconstructed, even though the waveform itself was compatible with the response to a single photon. Figure 8.4 shows an example of an ATWD readout with erroneous pulse reconstruction. The problem is caused by a too low ATWD baseline assumed in the delayed pulse extraction and is subject to current investigations within the IceCube collaboration[5].

To avoid the faulty pulse reconstruction in this work, the analysis is run on DOM launches instead of pulses. As the faint delayed signal is spread over a roughly 1 ms interval, the effect of ignoring photons that arrive within the time window of a single DOM launch is small. Figure 8.5 shows the echo detection efficiency for a launch-based

---

[5]Since this problem is a so far unknown detection artifact, it is not modeled in the full detector simulation. Results from pulse-based studies using the detector simulation datasets thus remain valid.
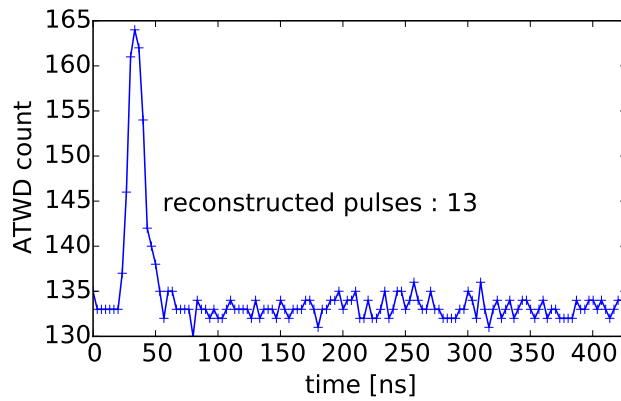
Figure 8.4: Example of an ATWD readout with faulty pulse reconstruction. Although the ATWD time spectrum has a single-pulse shape, 13 pulses were reconstructed from it.

analysis. No statistically significant difference in the detection efficiency is observable compared to the pulse-based approach (see Figure 7.6).
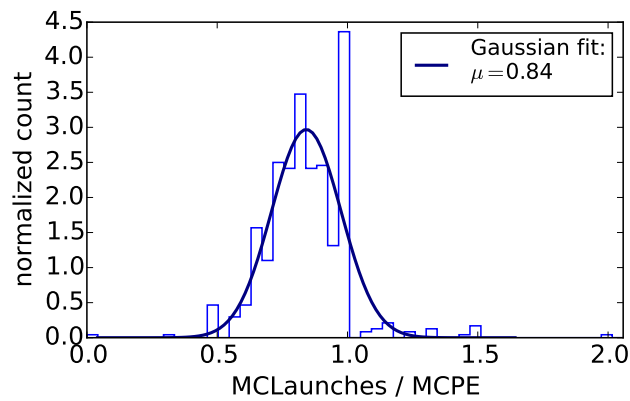

Figure 8.5: The number of delayed Monte Carlo simulated launches (MCLaunches) over the number of Monte Carlo delayed PE (MCPE) for around 600 showers at $100-150$ TeV energy. The peak at 1 is caused by the integer nature of the launch and MCPE count.

8.2.2   Results: individual event information

The delayed light deposition of 13 shower-like HESE events was recorded by the Hit-Spool system within the livetime of this analysis. The best-fit delayed hits, reconstructed under the neutron echo hypothesis, as well as basic information for all events contained in the analysis dataset are summarized in Table 8.1.

For the event with the strongest afterglow, recorded on June 26, 2017, the individual event information (page 100) and delayed launch spectrum (Figure 8.8) are presented in more detail in this section. The corresponding plots and information for the remaining showers in the analysis dataset can be found in Appendix E.

The cascade detected on June 26, 2017 has the second highest energy deposition in the data sample and a reconstructed event vertex at the center of IceCube's sensitive subdetector DeepCore. DeepCore is located in very clear ice, which provides ideal conditions for the detection of the faint echo. The likelihood fit attributes $45.5^{+2.4}_{-3.1}$ delayed launches to the neutron echo over a background expectation of 5.6 noise launches in the $30 - 1000\,\mu$s time window, which rejects the noise-only hypothesis at nearly $100\,\%$ confidence level[6]. Figure 8.6 and Table 8.2 furthermore indicate a large hadronic shower content. Respecting the $1\,\sigma$ systematic uncertainty range, Figure 8.7 (a) lets one conclude that at least around $70\,\%$ of similar ν_e CC showers would produce a fainter echo than the one observed. Figure 8.7 also demonstrates the impact of the large systematic uncertainty: within the $\pm 56\,\%$ uncertainty range, the value of the NC $\mathrm{PE_{delayed}}$-CDF at the best-fit $\mathrm{PE_{delayed}}$ (Figure 8.7 (c)) can nearly range from zero to one.
The eight DOMs included in the echo search almost equally contribute to the detection of the delayed SLC launches displayed in Figure 8.8. In addition, six HLC launches were recorded in two sets of two and three neighboring DOMs, respectively.

---

[6]using Wilk's theorem

event date	best-fit $\mathrm{PE_{delayed}}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2017/11/14	41	113 TeV	12 m	-385 m
2017/10/10	2	100 TeV	62 m	-375 m
2017/08/06	21	177 TeV	10 m	343 m
2017/06/26	46	397 TeV	29 m	-295 m
2017/05/11	11	118 TeV	23 m	-340 m
2017/03/23	1	82 TeV	49 m	-314 m
2017/03/04	13	202 TeV	42 m	-446 m
2016/12/26	4	165 TeV	45 m	185 m
2016/11/30	10	260 TeV	38 m	107 m
2016/11/06	4	65 TeV	39 m	50 m
2016/10/16	8	131 TeV	29 m	86 m
2016/10/06	22	132 TeV	38 m	-393 m
2016/07/11	12	622 TeV	56 m	242 m

Table 8.1: Delayed light deposition and event information of the 13 cascades in the analysis dataset. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
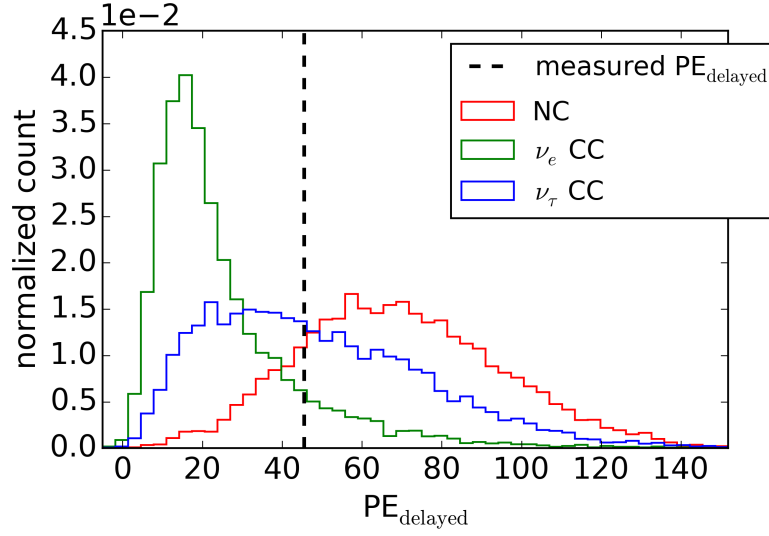
Figure 8.6: PE$_{\text{delayed}}$ distributions for the event hypothesis of the HESE shower recorded on June 26, 2017. The dashed line marks the best-fit PE$_{\text{delayed}}$.

	iCDF(PE$_{\text{delayed,fit}}$)
ν_{e} CC	0.13
ν_{τ} CC	0.51
NC	0.84

Table 8.2: Value of the inverse PE$_{\text{delayed}}$-CDFs at the best-fit PE$_{\text{delayed}}$ for the HESE shower recorded on June 26, 2017.

(a) ν_{e} CC   (b) ν_{τ} CC   (c) NC
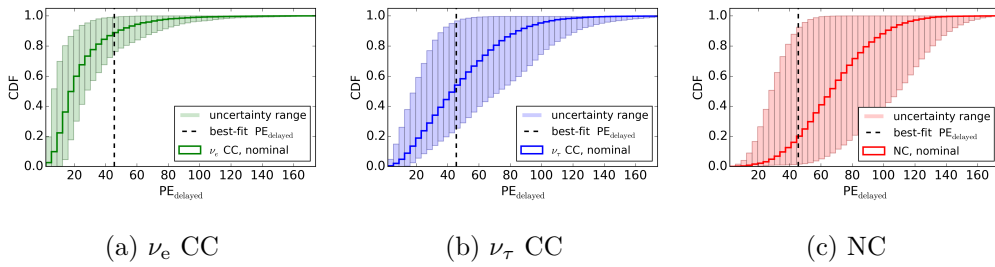
Figure 8.7: CDFs corresponding to the distributions in Figure 8.6. The bands show the $\pm 56\,\%$ systematic uncertainty range.
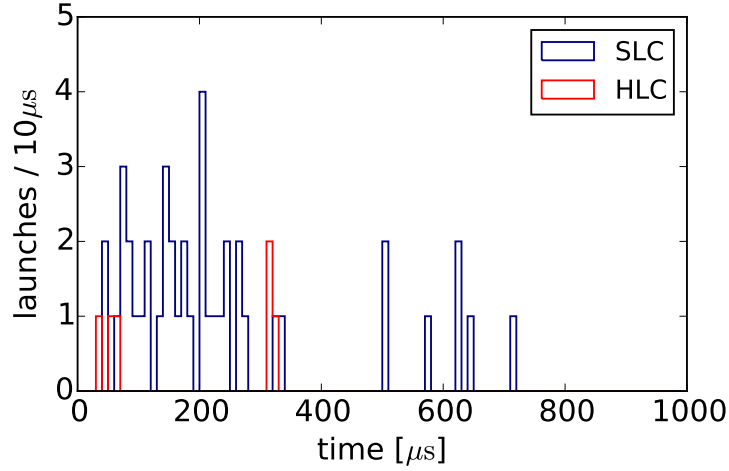
Figure 8.8: Delayed launch spectrum of the HESE shower recorded on June 26, 2017.

8.2.3 Results: ensemble analysis

In the analysis of the cascade ensemble, the likelihood and MCMC approach yield consistent results for the interaction type composition. Both methods find the 100 % ν_τ CC hypothesis describing the data best. The frequentist and MCMC triangle plot is shown in Figure 8.9 and Figure 8.10, respectively. The likelihood best-fit result for the nuisance parameter s is -0.05, which is compatible with the highest sample point of -0.13 in the MCMC. The MCMC corner plot for the analysis dataset is shown in Figure 8.11.

As indicated by the preceding studies, frequentist confidence regions in Figure 8.9 are rather conservative when compared to their Bayesian analog in Figure 8.10. However, a wide parameter space is allowed also within the Bayesian 68 % credible region, which again reflects the small sample size and large systematic uncertainty. The small sample size furthermore results in a wide minimum in the negative log-likelihood landscape which extends into the unphysical parameter range. As a consequence, the best-fit point tends to be located at the edge of the allowed parameter range defined in the likelihood boundary conditions.
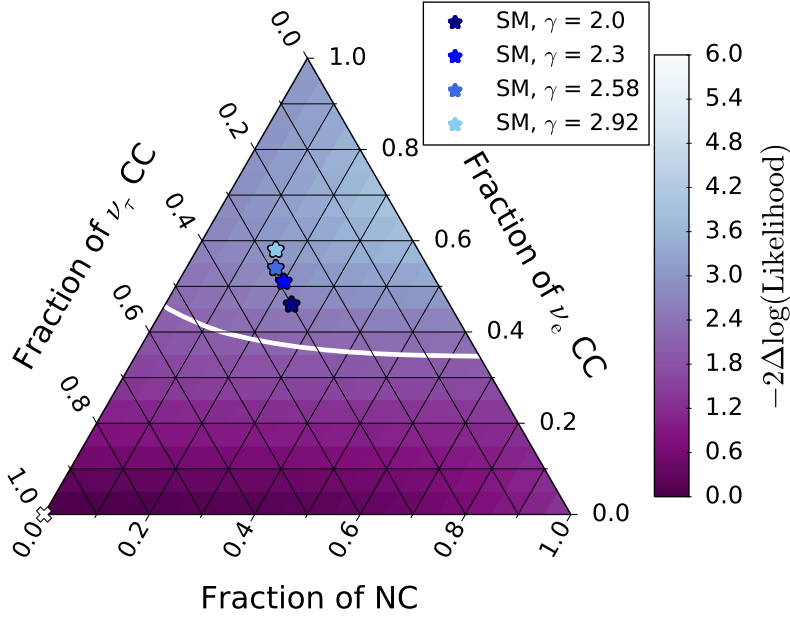
Figure 8.9: Two-dimensional likelihood scan around the best-fit point (white cross, [NC, ν_e CC, s] = [0.00, 0.00, −0.05]) of the analysis dataset. The solid and dashed white line refers to the 68 % and 95 % confidence interval, respectively. The 95 % confidence interval lies outside the displayed parameter space. Blue markers show the average shower type composition for the standard model (SM) hypothesis with different assumptions for the astrophysical index.

The *standard* assumption of an atmospheric and astrophysical neutrino flux as the origin of the observed showers is compatible with the data sample. The frequentist p-values and Bayesian observation probabilities for different assumptions for the astrophysical index (see Figure 8.9 and Figure 8.10) are given by:

assumed astrophysical index	p-value	observation probability
2	0.25	0.22
2.30	0.22	0.20
2.58	0.22	0.20
2.92	0.21	0.19

At last, the analysis dataset is used to probe the extreme 100 % NC and 100 % ν_e CC hypotheses, which are motivated in boosted dark matter theories (see Section 7.8.2). The wide frequentist confidence regions do not provide strong exclusion power, rejecting the 100 % NC and 100 % ν_e CC assumption at 55 % and 79 % confidence level, respectively. At 90 % confidence level, the one-dimensional likelihood scans shown in Figure 8.12 leave the maximal allowed NC fraction unconstrained and reject a relative ν_e CC contribution
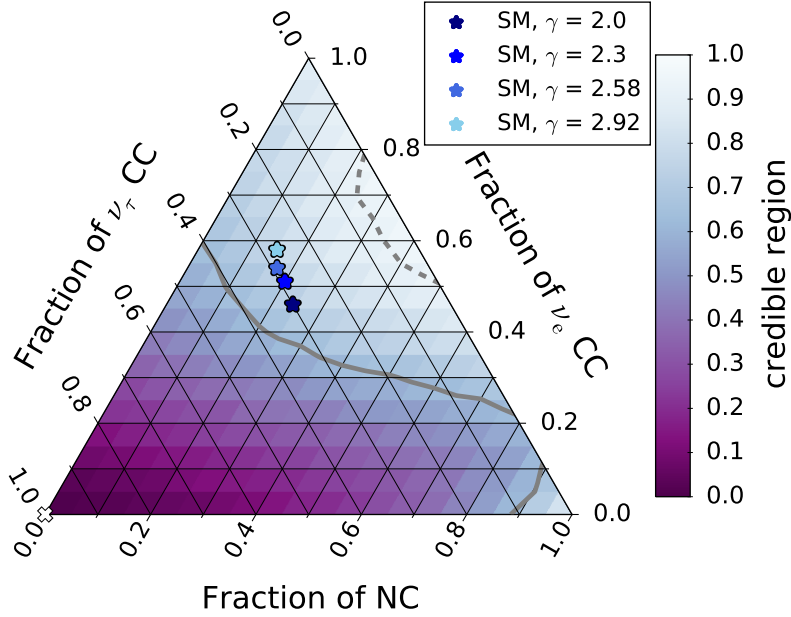
Figure 8.10: The triangle plot reflects the two-dimensional sample density of the MCMC for the analysis dataset. The solid and dashed gray line marks the 68 % and 95 % credible region, respectively. The highest density point ([NC, ν_e CC, ν_τ CC] = [0.00, 0.00, 1.00]) is shown as a white cross. Blue markers show the average shower type composition for the standard model (SM) hypothesis with different assumptions for the astrophysical index.

larger than 59 %. The narrower credible regions of the MCMC routine, however, assign an observation probability of less than 10 % to the parameter points associated with boosted dark matter scenarios. The point of 100 % NC and 100 % ν_e CC contribution is located at the edge of the 94 % and 91 % credible region, respectively.

Figure 8.13 shows the stacked spectrum of delayed launches in the analysis dataset, which demonstrates a clear delayed light signal well above the noise expectation. Using again an averaged deadtime correction function, the binned Poisson likelihood fit (see Section 8.1) finds an exponential decay time of $155^{+26}_{-21}\,\mu$s describing the excess best. A decay time of $217\,\mu$s, expected from Geant4 simulations, is consistent with the best-fit value at the $2\,\sigma$ level. The best-fit spectrum not shown in Figure 8.13 reveals the same problem as the best-fit spectrum of the low-PE dataset in Figure 8.2: the averaged deadtime correction function does not well reproduce the spectral features below $\mathcal{O}(100\mu s)$. Although the influence on the fitted number of delayed signal hits is small, a deeper understanding of DOM deadtime effects will be important to improve the echo analysis.
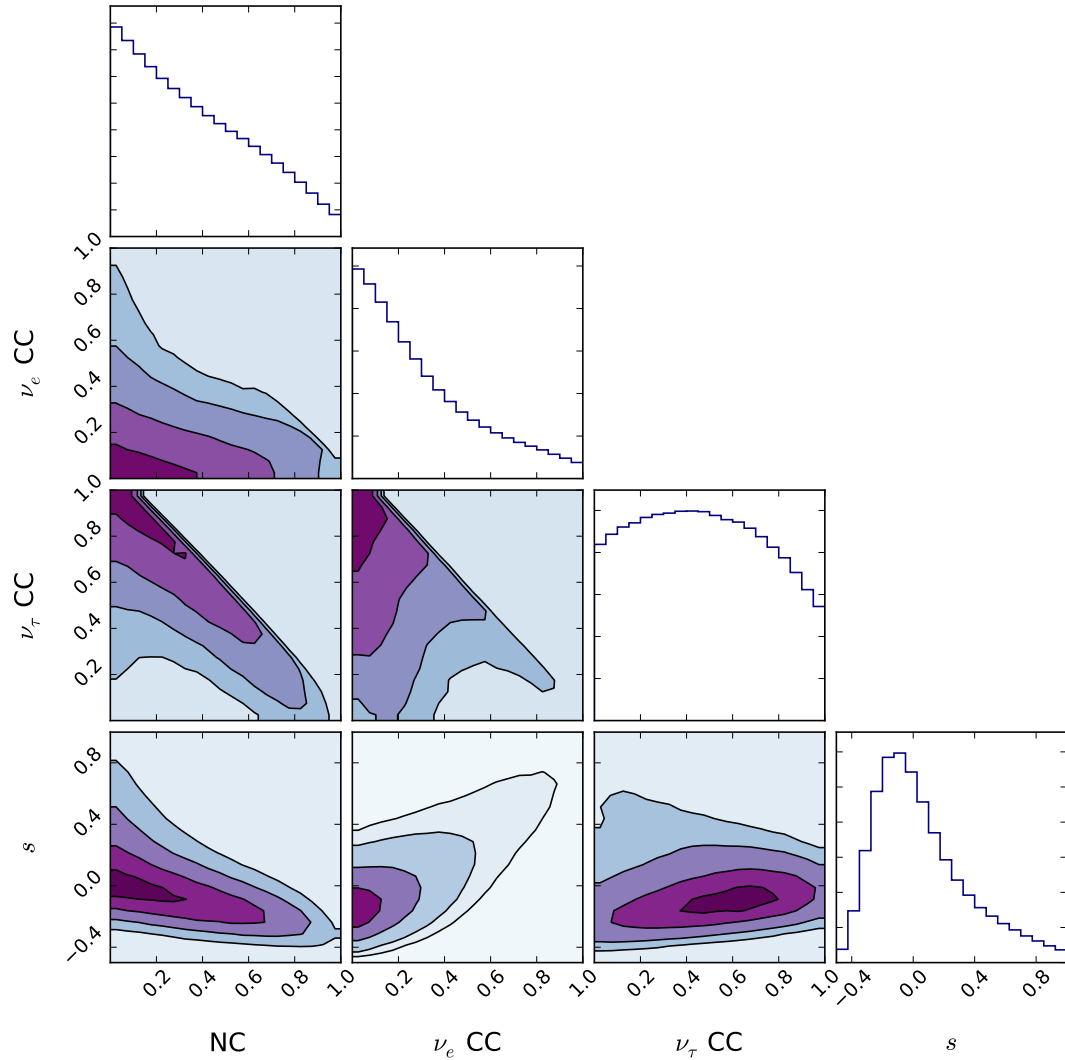
Figure 8.11: Corner plot of the MCMC result for the analysis dataset. The filled areas in the two-dimensional density plots refer to the $1\,\sigma$, $2\,\sigma$, $3\,\sigma$ and $4\,\sigma$ region. The boundary conditions $\alpha, \beta, \gamma \in [0,1]$ and $\alpha + \beta + \gamma = 1$ (see Equation 7.10 and Section 7.7.3) reject the upper right parameter space in the two-dimensional projections of the relative interaction type contributions. The allowed range for α, β and γ has been slightly extended to $[-0.1, 1.1]$ in the MCMC to avoid edge effects.

(a) NC

(b) ν_e CC
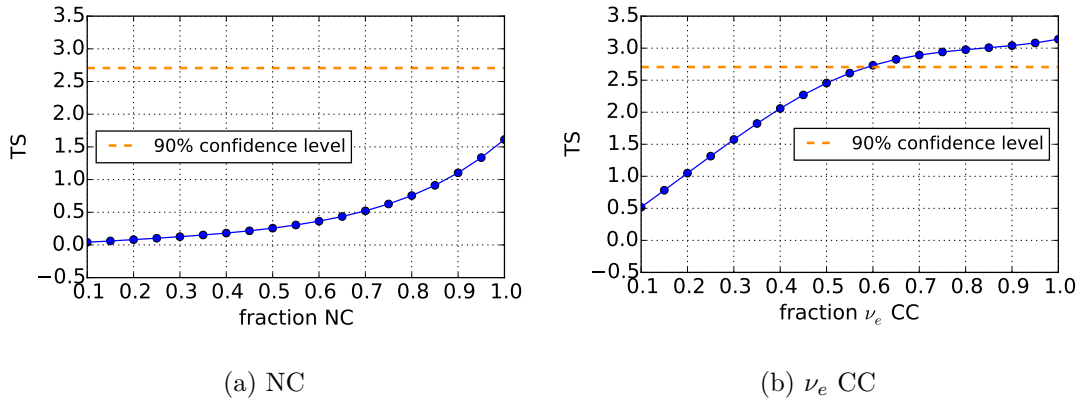
Figure 8.12: One-dimensonal likelihood scans for the analysis dataset. The minima of the displayed distributions extend into the range of negative fractions, which is not contained in the physically allowed parameter range.
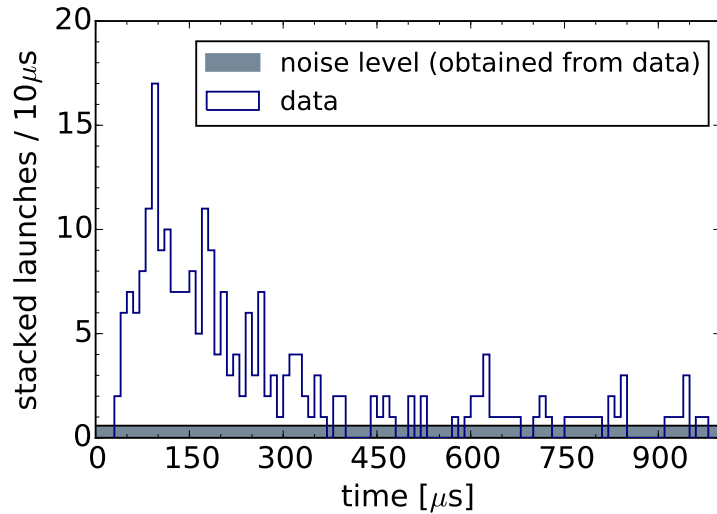
Figure 8.13: The stacked delayed launch spectrum of the analysis dataset. The noise expectation is directly extracted from the 0.5 s of data recorded before the actual event.

9
Summary and outlook

This work presents a method which, for the first time, allows to classify the interaction type of high energy cascades in IceCube: a general cascade type identification, not restricted to the search for the distinctive topology of ν_τ CC interactions, has never been implemented before. While the idea to exploit the neutron echo as an indicator of the hadronic shower content has been discussed by IceCube collaborators and others [25], in-depth studies and the experimental implementation, including the required alternative data acquisition, had not been attempted.

Extensive shower simulations, using the PYTHIA 8 and Geant4 toolkit as well as the full IceCube detector simulation chain, were performed to quantify the appearance of the echo in the detector and to determine the experimental requirements for its detection. Since a Geant4-based shower simulation turned out to be too CPU-intense and time consuming to be employed in the analysis, a routine that uses parametrized Geant4 results was developed and implemented. Furthermore, the uncertainties involved with an echo analysis in IceCube have been quantified and incorporated in the statistical evaluation. A likelihood approach to optimally extract the echo in the presence of a PMT dark rate was developed and a fast routine was implemented to evaluate the substantial DAQ-induced deadtime for delayed signals. A frequentist likelihood approach as well as a Markov Chain Monte Carlo algorithm is now in place to determine the interaction type composition of cascade samples.

The neutron echo arises around $100\,\mu$s after the neutrino interaction, which is beyond IceCube's standard readout windows. To capture the faint afterglow, a complete detector readout - so-called HitSpool data - is now requested for high energy events. Within the scope of this work, the automated request and processing of one second of HitSpool data for events identified by the *High Energy Starting Event* (HESE) filter was implemented. The HESE HitSpool system collects data since February 2016 in a stable fashion.

In addition to the analysis threshold of $6000\,$PE, HitSpool data is recorded for events in the $1500-6000\,$PE range. In studies of the low-PE sample, a clear excess at a delay time of around $100\,\mu$s in the photon count spectrum was observed. Although not investigated to the full extent, since atmospheric muons with catastrophic energy losses dominate the sample, this excess proved that an afterglow is visible in high energy events and strongly motivated the pursuit of the echo analysis.

Within the 22 months of livetime, 13 cascades with deposited energies from around 60 TeV to 600 TeV were recorded above the 6000 PE threshold. Individual shower type information is shown for every cascade in the data sample. The interaction type composition of the ensemble was analyzed with regard to its consistency with the *standard* hypothesis of pion- and kaon-induced astrophysical neutrino production for various assumptions for the spectral index of the astrophysical energy spectrum. The collected data are compatible with the hypothesis of an astrophysical neutrino flux as the main source of IceCube's high energy events. Although the so far acquired statistics is low, the hypothesis of an either solely hadronic or electromagnetic shower type composition, which is motivated by boosted dark matter theories, is already disfavored with a Bayesian observation probability of less than 10 %.

The cascade type identification method was applied in this work to search in the HESE shower sample for hints of new physics. It will, however, also greatly advance non-exotic analyses of the astrophysical neutrino signal. The routines developed in the course of this thesis will be incorporated in IceCube's global flavor fit presented in Section 4.2, whose power suffered from the indistinguishability of the shower-like electron and tau neutrino event signature. When applied in a Standard Model analysis, the neutral current fraction, which can conservatively be pinned to the $15 - 25\,\%$ interval, can be constrained in the ensemble fit, thereby enhancing the discrimination power for the electron and tau flavor.

Furthermore, the identification of Glashow resonance events will strongly profit from the novel analysis. Attributing a large hadronic shower content to a cascade at the 6 PeV scale would substantiate its link to a Glashow resonance interaction. If a shower with the characteristic energy deposition is detected in the near future, the computationally expensive full detector simulation, including Geant4 particle propagation, will be run to produce the required delayed photon distributions with minimal systematic uncertainty.

The reduction of the involved systematic uncertainties is essential in the further improvement of the neutron echo analysis. The fast routine calculating the DOM deadtime effects should be revised, the basis of Geant4 simulated events should be enhanced and approximations should be replaced by more advanced descriptions. One can also think of approaches other than the parametrization of Geant4 results to substitute the CPU-expensive Geant4 simulation in the full detector Monte Carlo.

Furthermore, experiments are under preparation whose results will reduce the uncertainty involved with the modeling of hadronic processes in general and the neutron echo in particular. The proposed *Accelerator Neutrino Neutron Interaction Experiment* (ANNIE) [110], which will measure the neutron yield of neutrinos interacting in gadolinium-doped water at atmospheric neutrino energies, will provide the data necessary to cross-validate and improve the hadronic models implemented in the simulation

tools used here. The IceCube collaboration plans to deploy detectors in the 1750 m deep liquid-filled *SPICEcore* hole to measure the neutron capture time profile at IceCube conditions as well as the luminescence of ice [65], exploring luminescence as an alternative source of delayed light.

Recent studies [111] showed that the accuracy, at which the method is able to determine the interaction type composition of a data sample, depends stronger on the sample size than on the assumed systematic uncertainty. The analysis' potential will thus increase with the number of high energy cascades recorded in the HESE HitSpool stream. In the same way, the routine would be extremely powerful for the high statistics datasets recorded with the proposed *IceCube-Gen2* high energy array [112]. Although the design is not yet finalized, *IceCube-Gen2* will extend IceCube to cover a surface area of around 10 km^2 and reach even lower depths. Since the ice in the target areal is extremely clear and more advanced optical modules will be deployed, the inter-string spacing can be extended to approximately 250 m without losing in detection efficiency or resolution for high energy events. The current limitation of the DAQ, which introduces up to 40 % deadtime for delayed signals, will not be present in the *IceCube-Gen2* design. The event rate of astrophysical neutrinos is expected to be ten times larger in *IceCube-Gen2* than in the current detector configuration. Since the echo analysis gains from small vertex-to-string distances, the number of cascades for which the analysis can be run will, however, not increase at the same magnitude as the event rate. Assuming a four times higher event rate for the echo analysis, a two times higher detection efficiency for Cherenkov photons, no DAQ deadtime for delayed signals and implementing a Gaussian constraint for the NC fraction ($(20 \pm 5)\%$), Figure 9.1 demonstrates the wide potential of the neutron echo analysis in *IceCube-Gen2* and in IceCube's future flavor analyses.
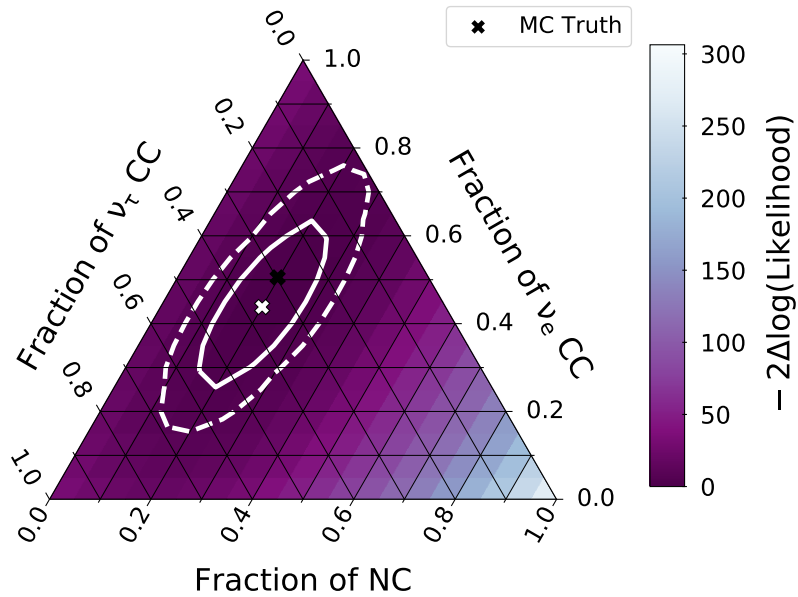
Figure 9.1: The triangle plot was produced within the scope of a Bachelor's thesis [111]. The input data for the two-dimensional likelihood scan was a dataset consisting of around 1000 simulated showers, re-weighted to a sample size of 84 events, which is a first estimate for the sample size for three years of *IceCube-Gen2* data. The solid and white dashed line refers to the 68 % and 95 % confidence region, respectively. In the likelihood function, the NC fraction was constrained via a Gaussian penalty factor with mean 20 % and standard deviation 5 %. When run on a larger data sample and with a prior for the NC contribution, the neutron echo analysis shows great potential.

Appendices

A

GEANT4 physics lists

The following appendix gives a brief description of the physics lists used in this analysis.

A.1 Reference physics lists

The Reference Physics Lists are named according to the following acronyms referring to the implemented hadronic models, which are described in detail in the Geant4 Physics Reference Manual [83]:

- **QGSP:** Quark-Gluon String Precompound model for high energy hadronic collisions. The Quark-Gluon String model describes the hadron-nucleus interaction including hadronization, while the Precompound part models the subsequent de-excitation of the nucleus.

- **FTFP:** FRITIOF Precompound model for high energy hadronic collisions. The FRITIOF part describes the hadron-nucleus interaction, while the hadronization is described by the Lund fragmentation model. The Precompound part models the subsequent de-excitation of the nucleus.

- **BERT:** BERTini model for hadron inelastic scattering for primary protons, neutrons, pions and kaons below around 10 GeV.

- **BIC:** BInary Cascade model for hadron inelastic scattering for primary protons, neutrons, pions and kaons below around 10 GeV.

- **HP:** High Precision neutron model for neutron interactions below 20 MeV.

Both, the *QGSP_BERT_HP* and *QGSP_BIC_HP* physics list, are recommended by the Geant4 user support pages [113] for the study of LHC neutron fluxes and linear collider neutron fluxes. In this work, the simulation of the Cherenkov effect was implemented in addition.

A.2 Hadr06-based physics list

The physics list used in early Geant4 stand-alone simulations was based on the build-in example *Hadr06*, which is supposed to demonstrate the energy deposition and particle

flux in a hadronic cascade. The modified physics list, which was used as an alternative in this analysis, contains the following modules:

- **G4HadronPhysicsFTFP_BERT_HP:** FRITIOF Precompound model for high energy hadronic collisions and Bertini model for hadron inelastic scattering for primary protons, neutrons, pions and kaons below around 10 GeV. Additionally, the High Precision neutron model is included.

- **HadronElasticPhysicsHP:** This module registers G4NeutronHPThermalScattering for neutrons below 4 eV in addition to the standard elastic scattering of the High Precision neutron model. It thus takes atomic translational motion as well as atomic vibration and rotation, relevant at these low energies, into account. However, since this model is data driven, thermal scattering data for the implemented detector material must be available in order to register the thermal scattering model. Unfortunately, no data was available for the detector material of this analysis.

- **EmStandardPhysics:** This is a variant of G4EmStandardPhysics which implements all standard electromagnetic processes in Reference Physics lists like *QGSP_BERT_HP*.

- **GammaPhysics:** This module handles photonuclear reactions. Since photons interact with the nucleus similar to hadrons, the interaction is modeled by a Bertini-style cascade below around 10 GeV and by the QGSP model in the high energy regime.

- **LeptoNuclearPhysics:** This module handles leptonuclear reactions for electrons, positrons and muons trough the corresponding models G4ElectroVDNuclearModel and G4MuonVDNuclearModel.

- **G4IonPhysics**

- **G4DecayPhysics**

- **G4RadioactiveDecayPhysics**

In this work, the simulation of the Cherenkov effect was implemented in addition.

This chapter holds the histograms and fit results of the *Fast Simulation's* photon count parametrization (see Section 7.2). Note that the displayed energy range may differ between the histograms. Some distributions show large statistical fluctuations that cannot be smoothed by the KDE smoothing algorithm. However, given the minor contribution of n, $\overline{\text{n}}$, p, $\overline{\text{p}}$, $\mu^{\pm}$ and e$^{\pm}$ to the total delayed photon yield, these uncertainties are less of a concern.
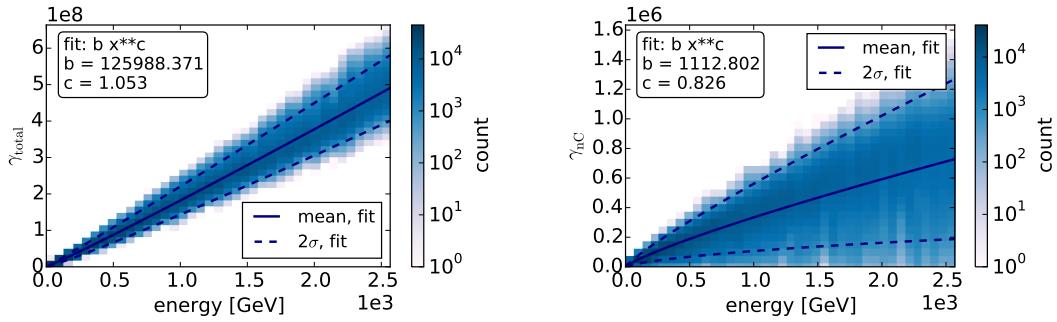
Figure B.1: Photon count parametrization for π^{+}.
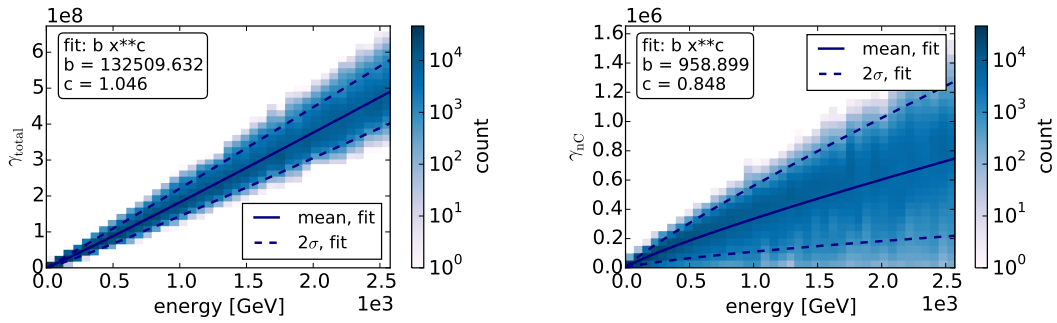
Figure B.2: Photon count parametrization for π^{-}.
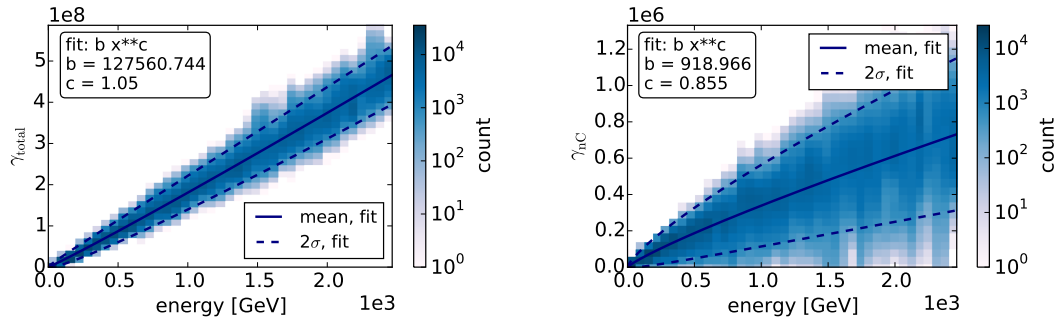
Figure B.3: Photon count parametrization for K$^-$.
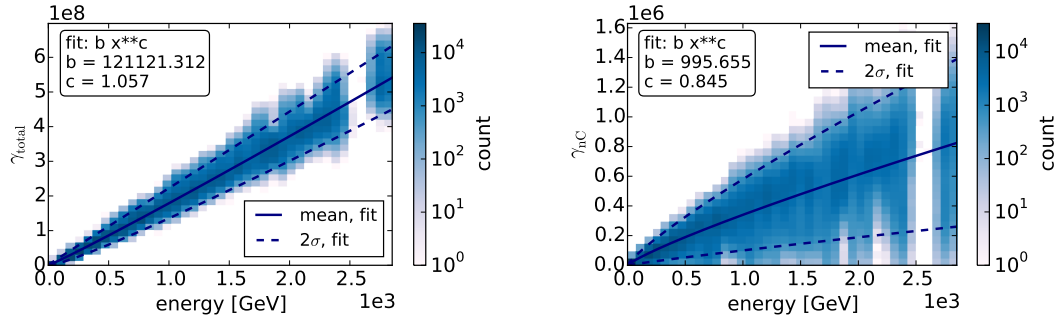
Figure B.4: Photon count parametrization for K^0_L.
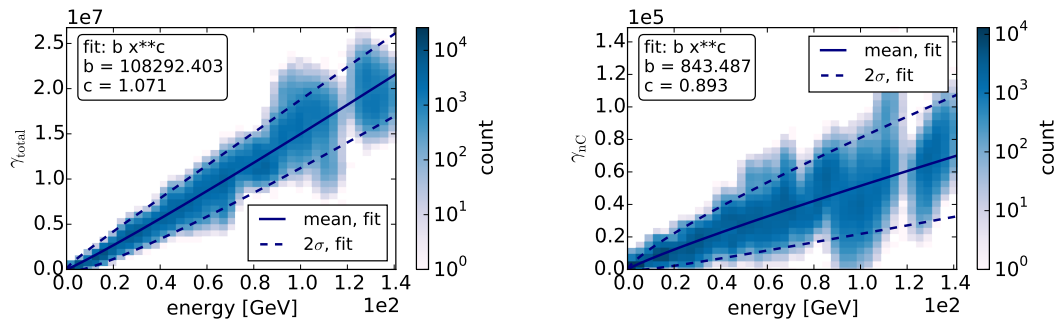
Figure B.5: Photon count parametrization for n.
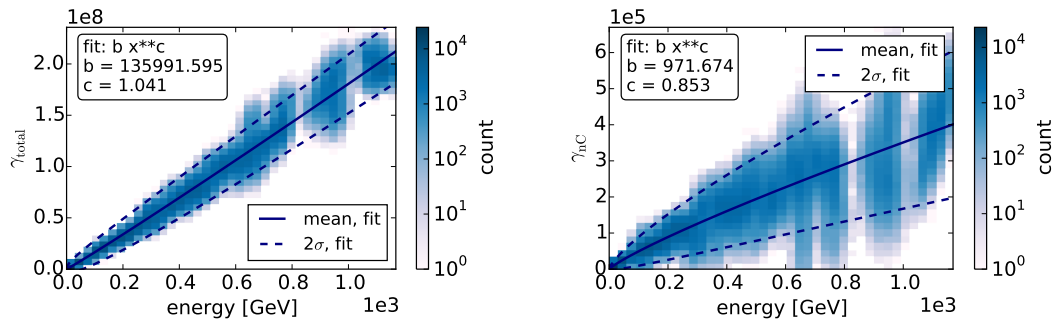
Figure B.6: Photon count parametrization for $\overline{\mathrm{n}}$.
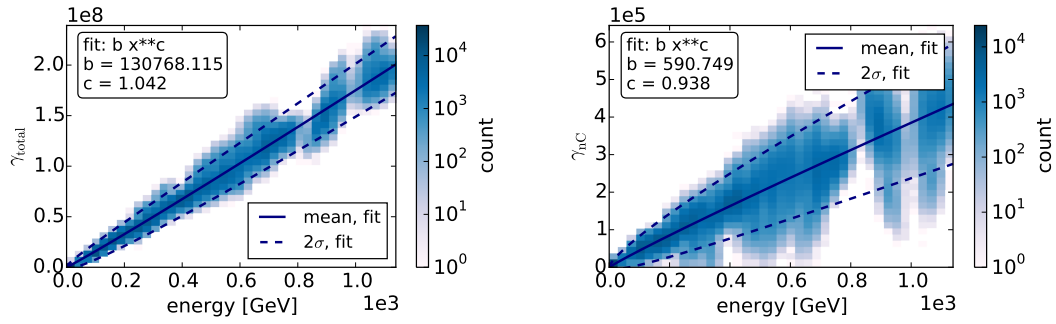
Figure B.7: Photon count parametrization for p.
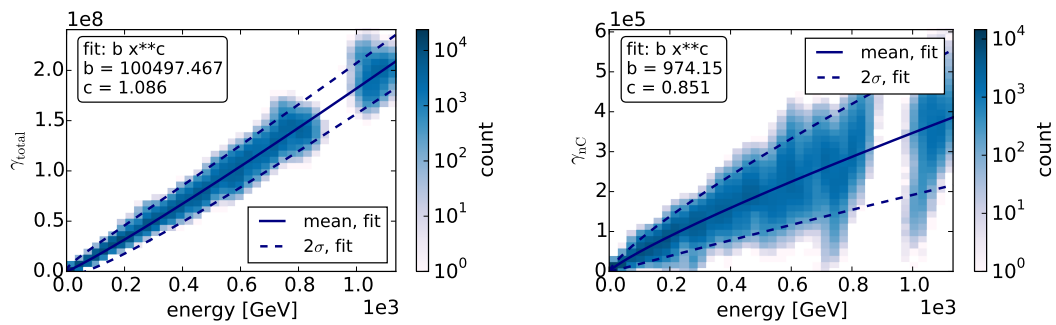
Figure B.8: Photon count parametrization for $\overline{\mathrm{p}}$.
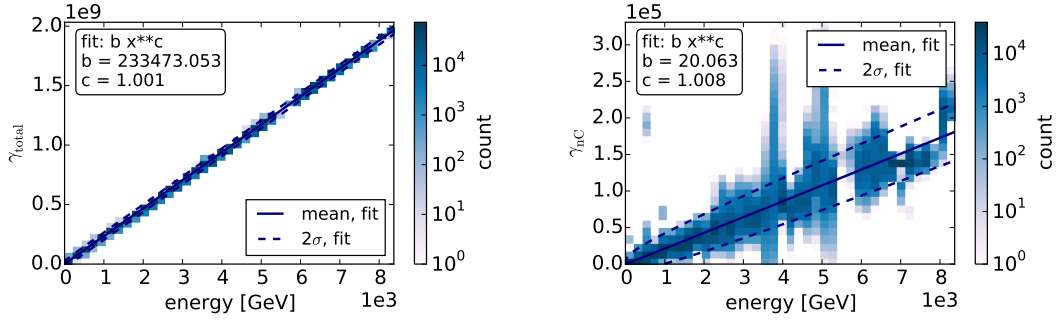
Figure B.9: Photon count parametrization for e$^+$ and e$^-$. Particle and antiparticle were combined to increase the statistics for the KDE and fit.
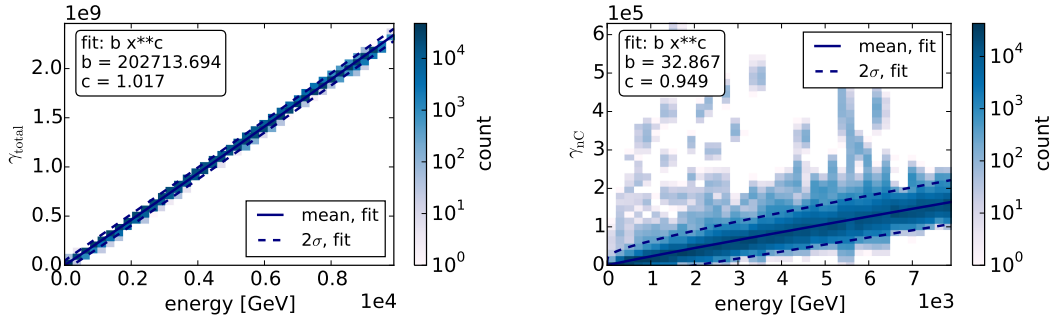
Figure B.10: Photon count parametrization for γ.
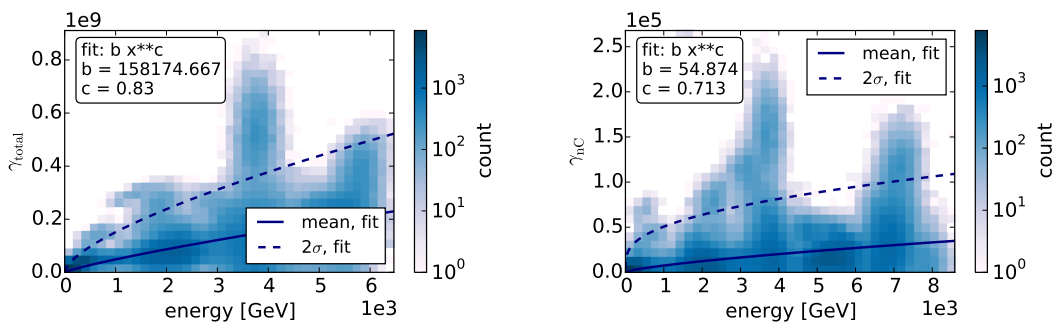
Figure B.11: Photon count parametrization for μ^+ and μ^-. Particle and antiparticle were combined to increase the statistics for the KDE and fit.

C Energy distribution in $100 - 150\,\mathrm{TeV}$ hadronic showers

To asses the typical energy of the most relevant particles in $\mathcal{O}(100\,\mathrm{TeV})$ hadronic showers, the following histograms show the CDF of the particle energy distributions with the ninetieth percentile marked.
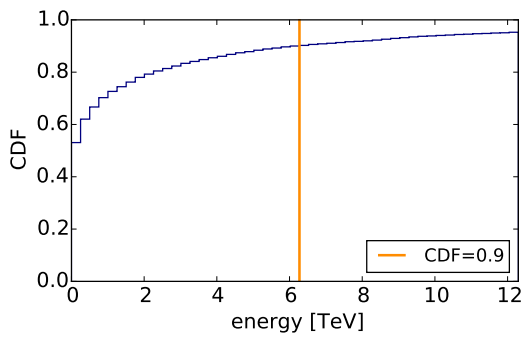
Figure C.1: π^+
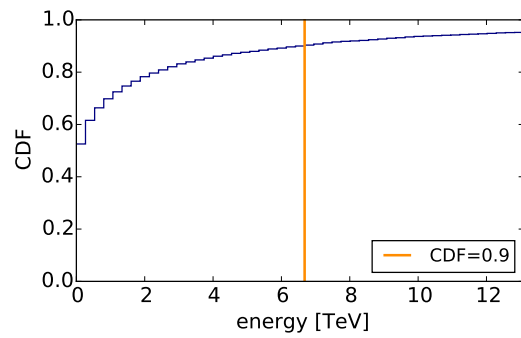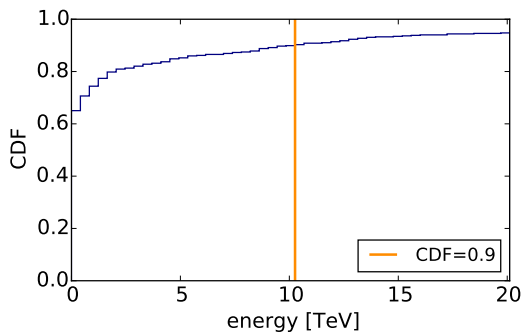
Figure C.2: π^-
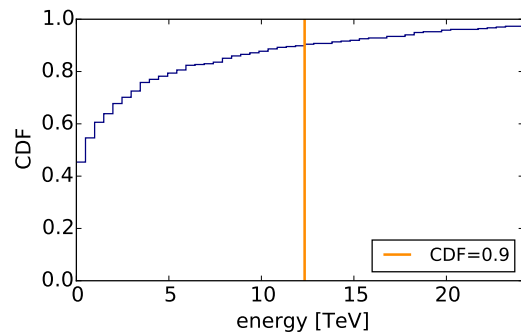
Figure C.3: n

Figure C.4: $\overline{\mathrm{n}}$
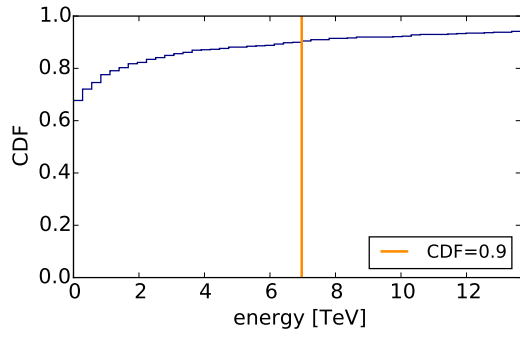
Figure C.5: p
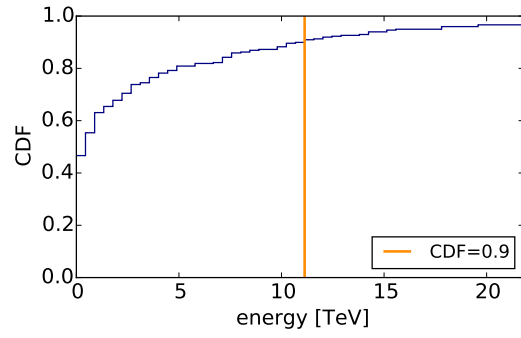
Figure C.6: $\overline{\text{p}}$
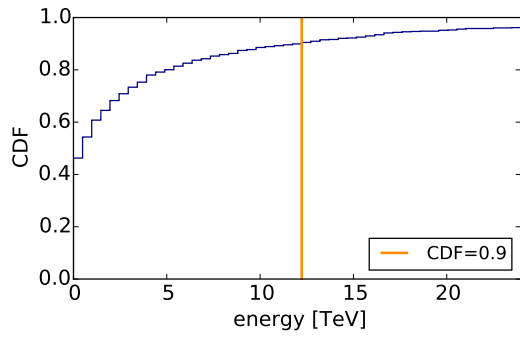
Figure C.7: K^+
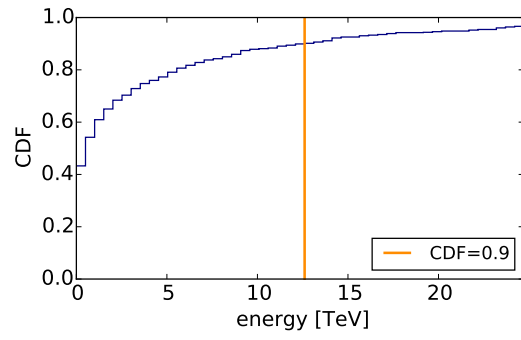
Figure C.8: K^-
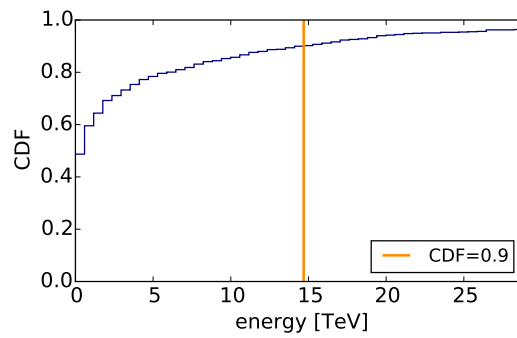
Figure C.9: K_L^0

Full detector simulation and *Fast Simulation* for a realistic simulation dataset

In order to validate the consistency of the full detector simulation and the *Fast Simulation*, both simulation chains were run for 13 simulated events, which represent a realistic analysis dataset. The following compilations show three figures for each simulated shower: at the top an event view, at the lower left the *Fast Simulation* result for the shower's true particle content and at the bottom right the *Fast Simulation* result for an unknown shower content. The bottom row plots additionally display the best-fit $PE_{delayed}$ value for the full detector simulation. The interaction type and the event's reconstructed energy are given in the caption.
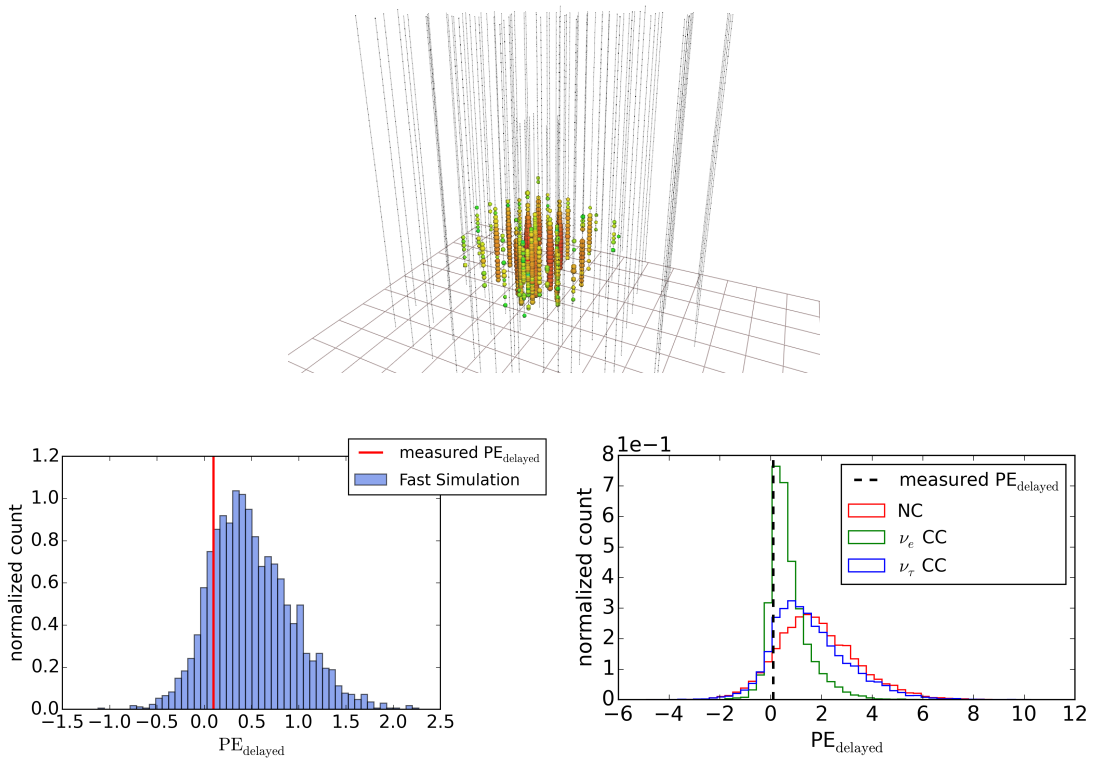
Figure D.1: Simulated ν_e CC shower, 67 TeV reconstructed energy.
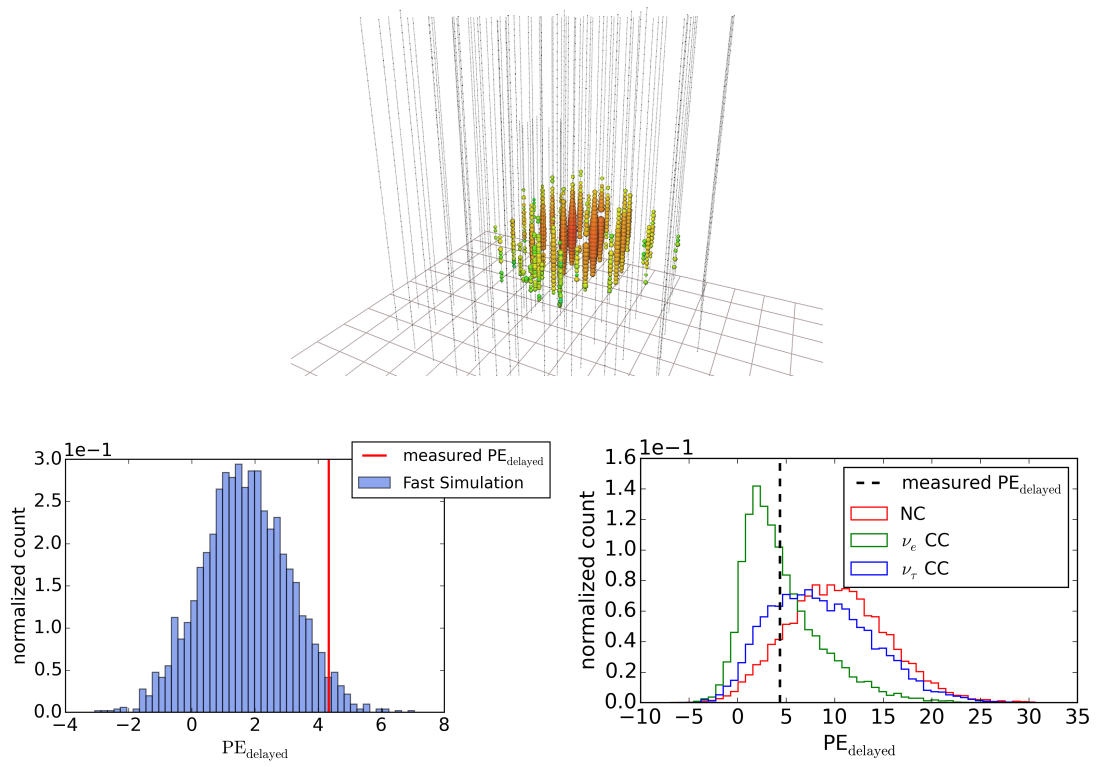
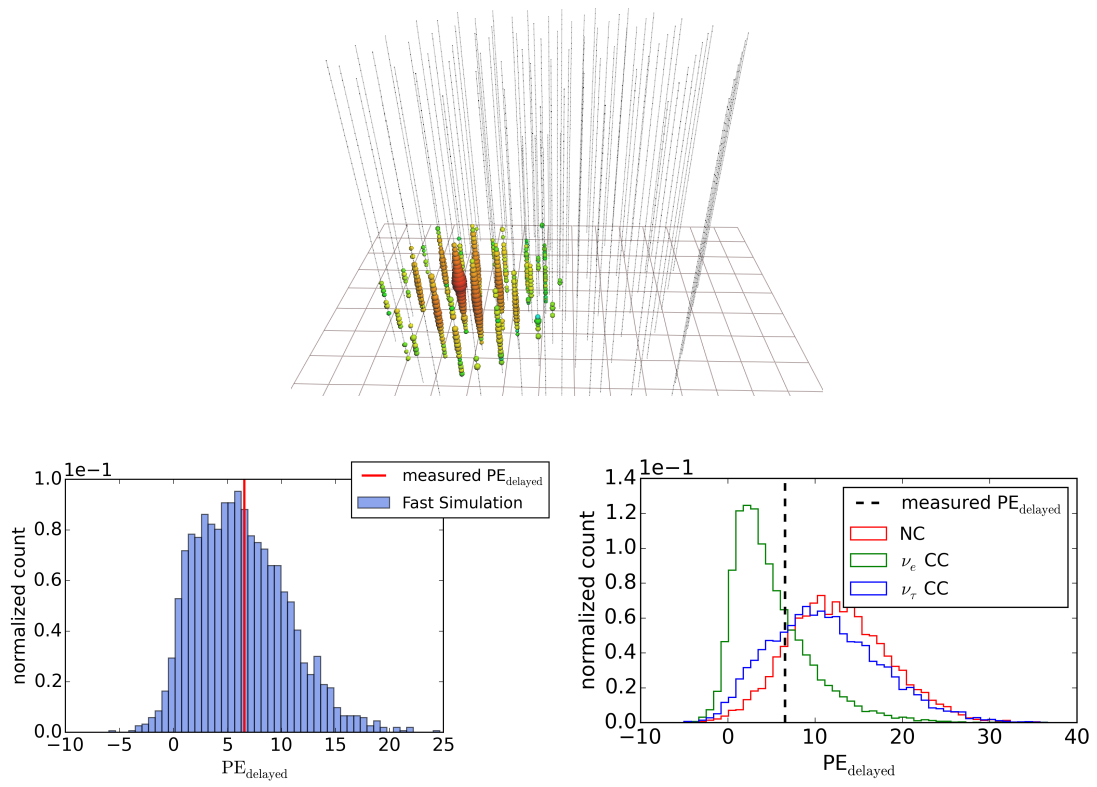Figure D.2: Simulated ν_e CC shower, 141 TeV reconstructed energy.

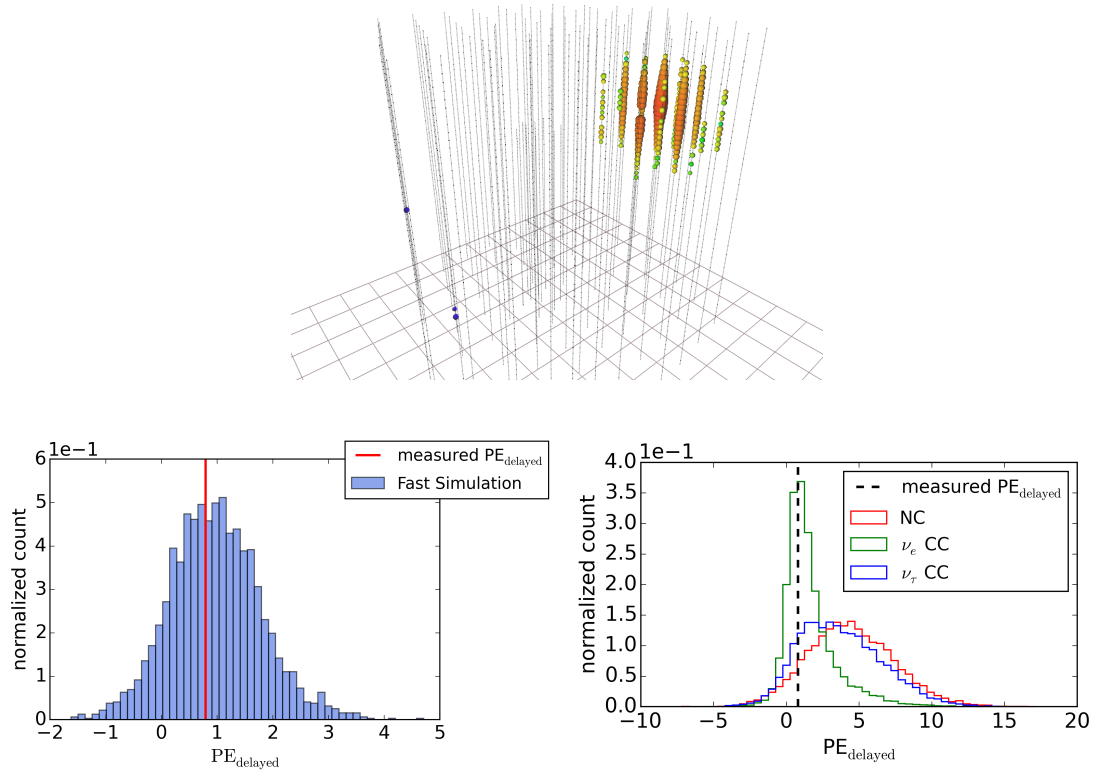Figure D.3: Simulated ν_e CC shower, 64 TeV reconstructed energy.

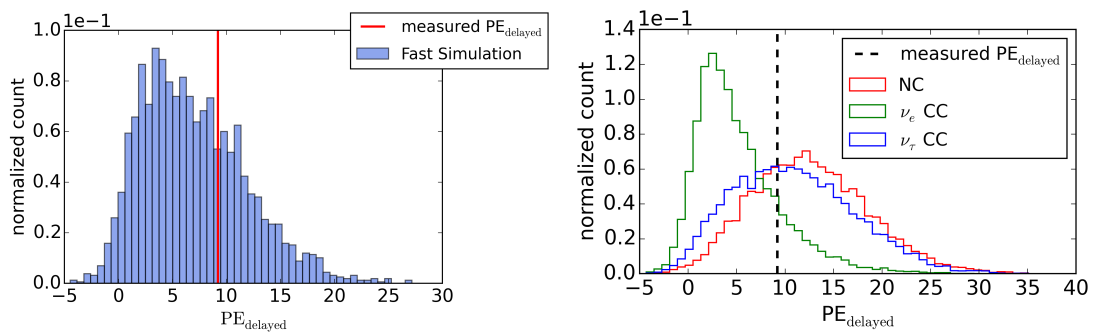Figure D.4: Simulated ν_e CC shower, 237 TeV reconstructed energy.

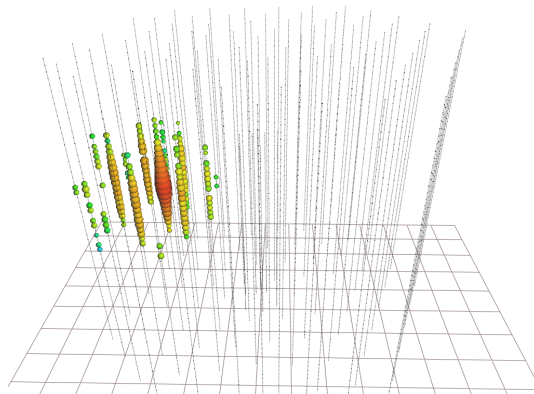Figure D.5: Simulated ν_e CC shower, 106 TeV reconstructed energy.

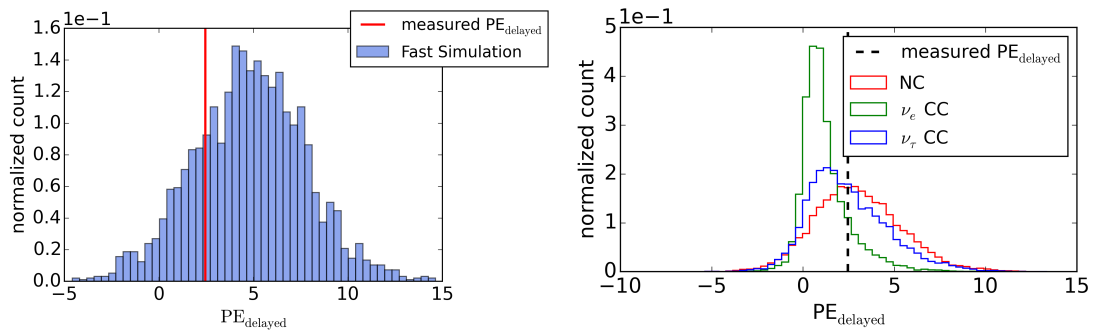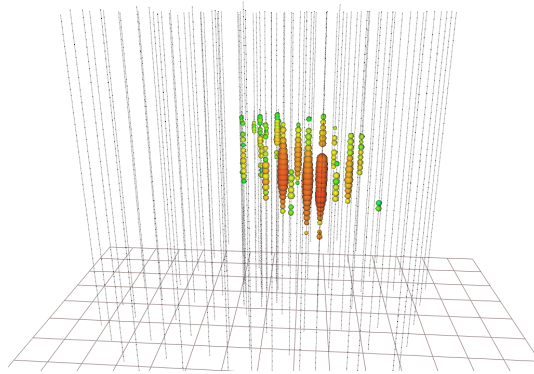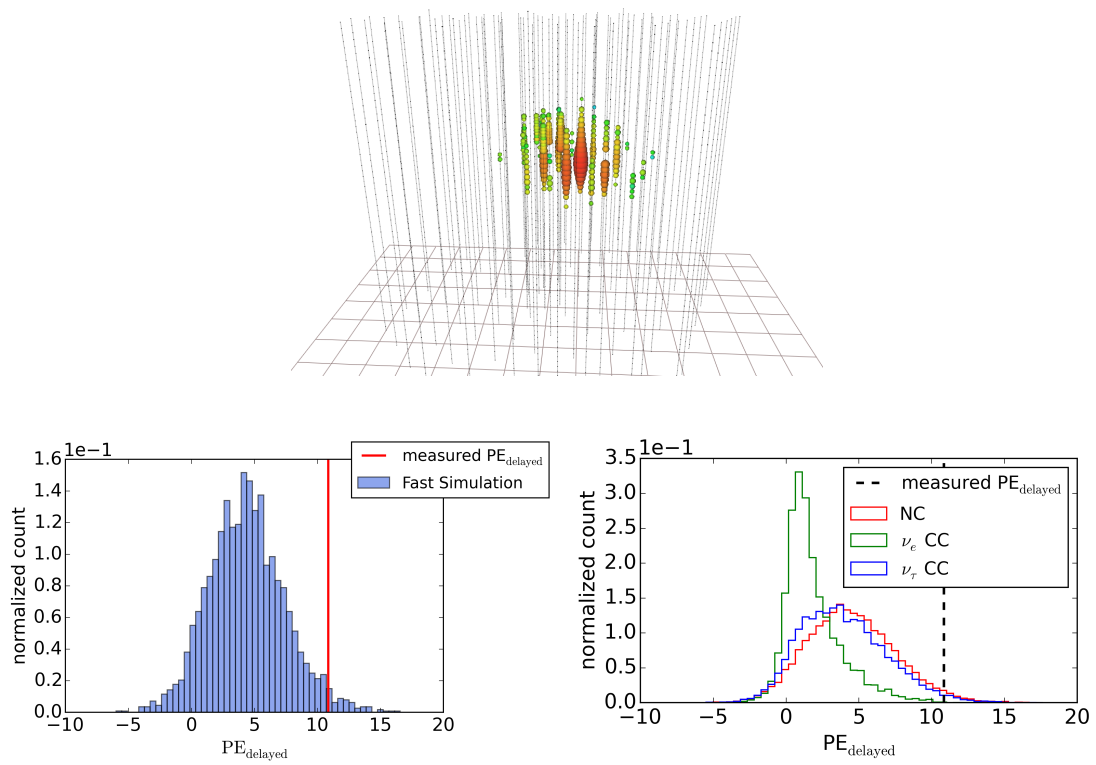Figure D.6: Simulated ν_τ CC shower, 734 TeV reconstructed energy.

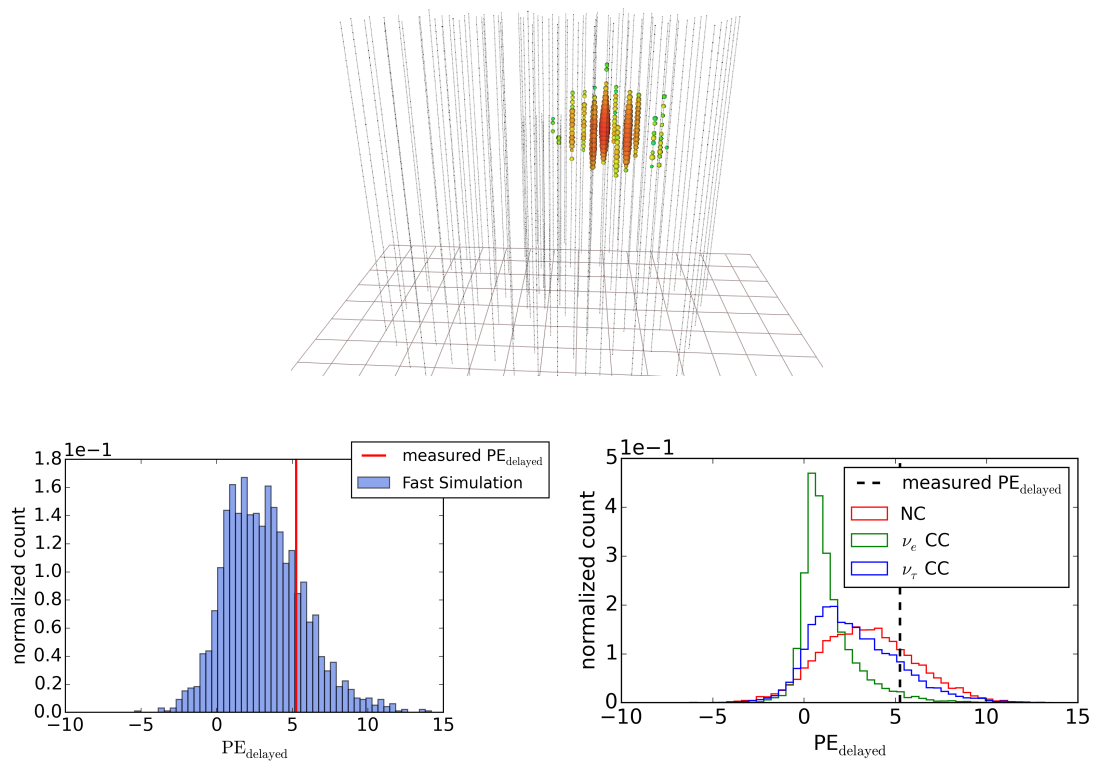Figure D.7: Simulated ν_τ CC shower, 101 TeV reconstructed energy.

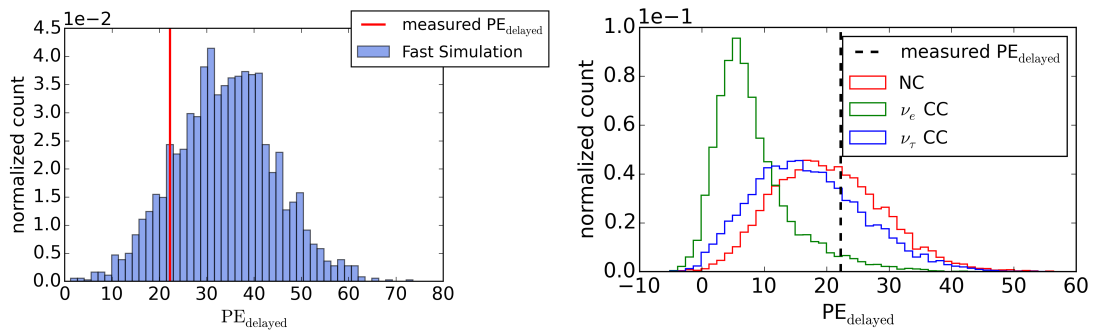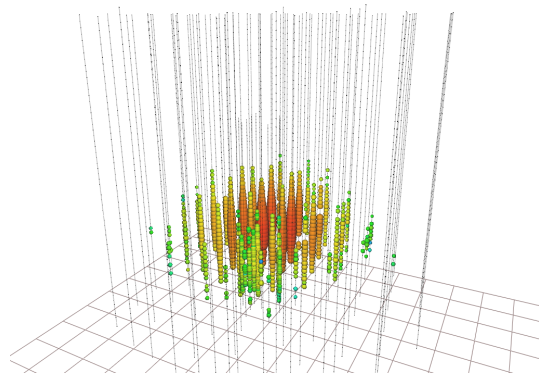Figure D.8: Simulated ν_τ CC shower, 105 TeV reconstructed energy.

Figure D.9: Simulated ν_e NC shower, 787 TeV reconstructed energy.

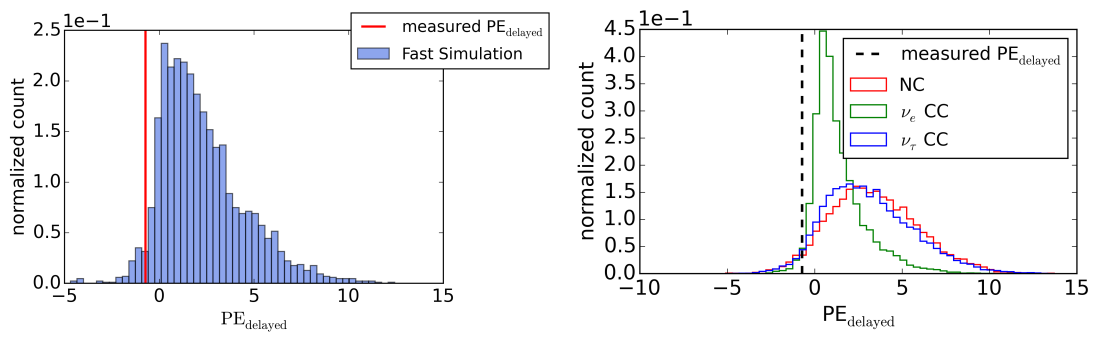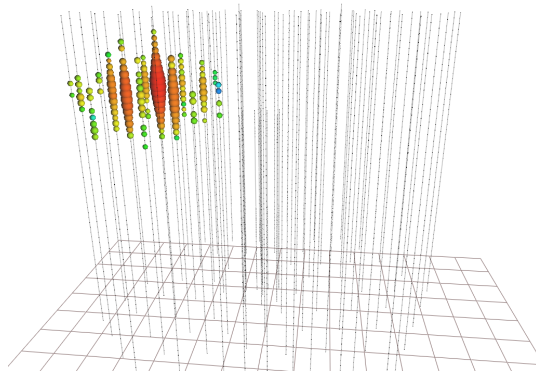Figure D.10: Simulated ν_μ NC shower, 77 TeV reconstructed energy.
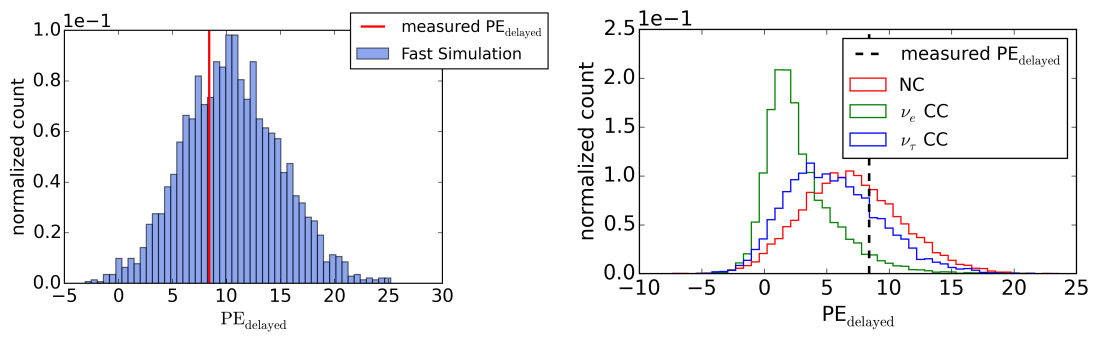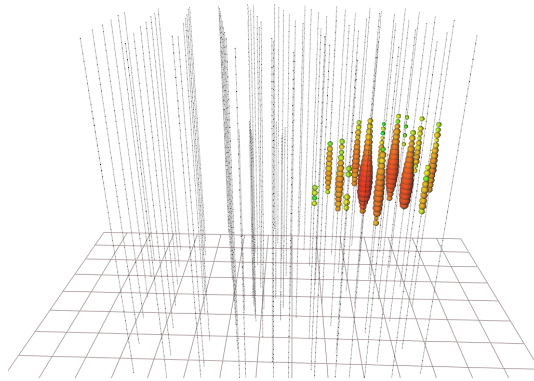
Figure D.11: Simulated ν_μ NC shower, 374 TeV reconstructed energy.
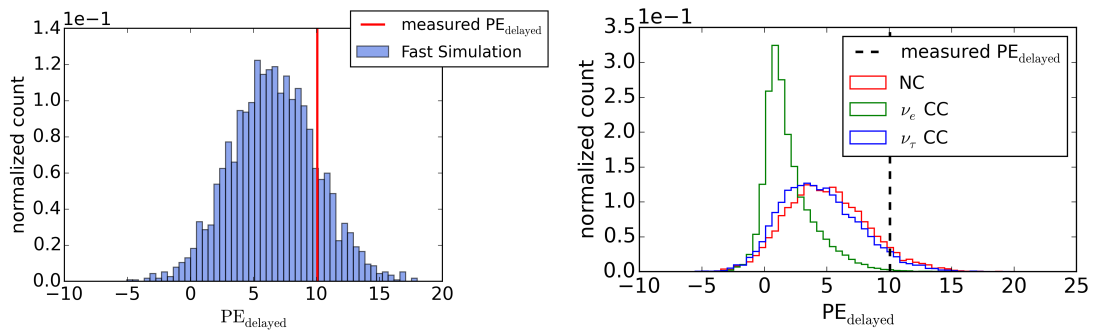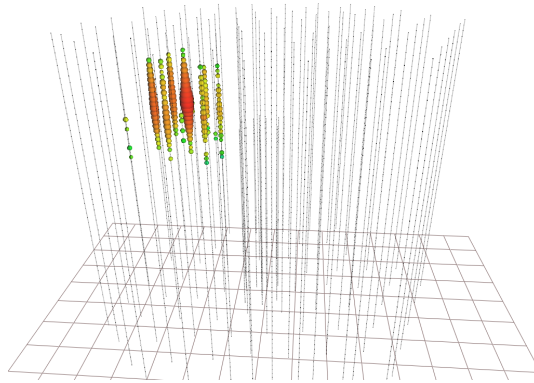
Figure D.12: Simulated ν_τ NC shower, 117 TeV reconstructed energy.

E

Individual event information for all showers in the analysis dataset

This chapter contains the best-fit $PE_{delayed}$ values, $PE_{delayed}$ probability density distributions produced by the *Fast Simulation* and event information for the 13 showers in the analysis dataset. For the sake of completeness, the event information of the shower already presented in Section 8.2 is also included.

Event 1

event date	best-fit $PE_{delayed}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2017/11/14	41	113 TeV	12 m	-385 m

Table E.1: Delayed light deposition and event information of the HESE shower recorded on November 14, 2017. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.

Figure E.1: $PE_{delayed}$ distributions for the event hypothesis of the HESE shower recorded on November 14, 2017. The dashed line marks the best-fit $PE_{delayed}$.

	iCDF(PE$_{\text{delayed,fit}}$)
ν_e CC	0.04
ν_τ CC	0.24
NC	0.46

Table E.2: Value of the inverse PE$_{\text{delayed}}$-CDFs at the best-fit PE$_{\text{delayed}}$ for the HESE shower recorded on November 14, 2017.

(a) ν_e CC   (b) ν_τ CC   (c) NC

Figure E.2: CDFs corresponding to the distributions in Figure E.1. The bands show the $\pm 56\,\%$ systematic uncertainty range.

Figure E.3: Delayed launch spectrum of the HESE shower recorded on November 14, 2017. 4.56 noise launches are expected in the $30 - 1000\,\mu$s time window.

Event 2

event date	best-fit $\text{PE}_{\text{delayed}}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2017/10/10	2	100 TeV	62 m	-375 m

Table E.3: Delayed light deposition and event information of the HESE shower recorded on October 10, 2017. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
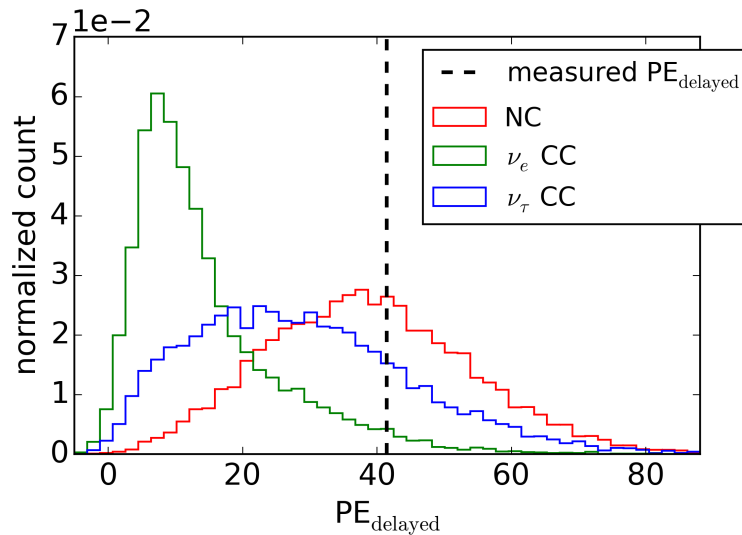

Figure E.4: $\text{PE}_{\text{delayed}}$ distributions for the event hypothesis of the HESE shower recorded on October 10, 2017. The dashed line marks the best-fit $\text{PE}_{\text{delayed}}$.

	iCDF(PE$_{\text{delayed,fit}}$)
ν_{e} CC	0.43
ν_{τ} CC	0.72
NC	0.86

Table E.4: Value of the inverse PE$_{\text{delayed}}$-CDFs at the best-fit PE$_{\text{delayed}}$ for the HESE shower recorded on October 10, 2017.

(a) ν_{e} CC
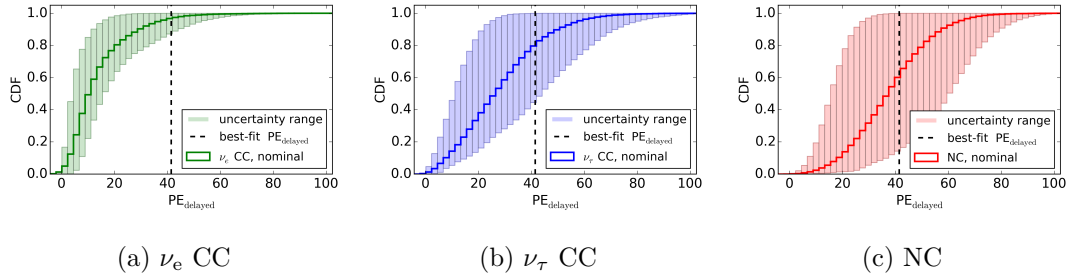
(b) ν_{τ} CC

(c) NC

Figure E.5: CDFs corresponding to the distributions in Figure E.4. The bands show the $\pm 56\,\%$ systematic uncertainty range.
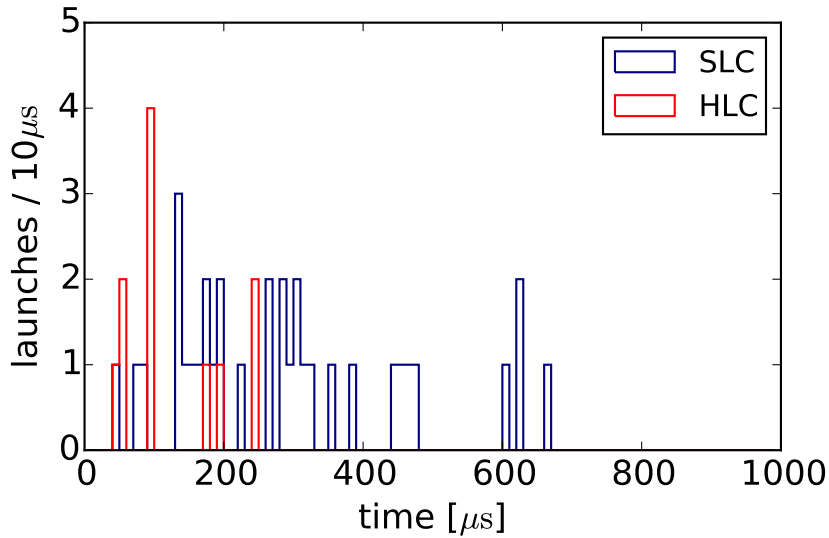

Figure E.6: Delayed launch spectrum of the HESE shower recorded on October 10, 2017. 3.75 noise launches are expected in the $30 - 1000\,\mu$s time window.

Event 3

event date	best-fit $\text{PE}_{\text{delayed}}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2017/08/06	21	177 TeV	10 m	343 m

Table E.5: Delayed light deposition and event information of the HESE shower recorded on August 6, 2017. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
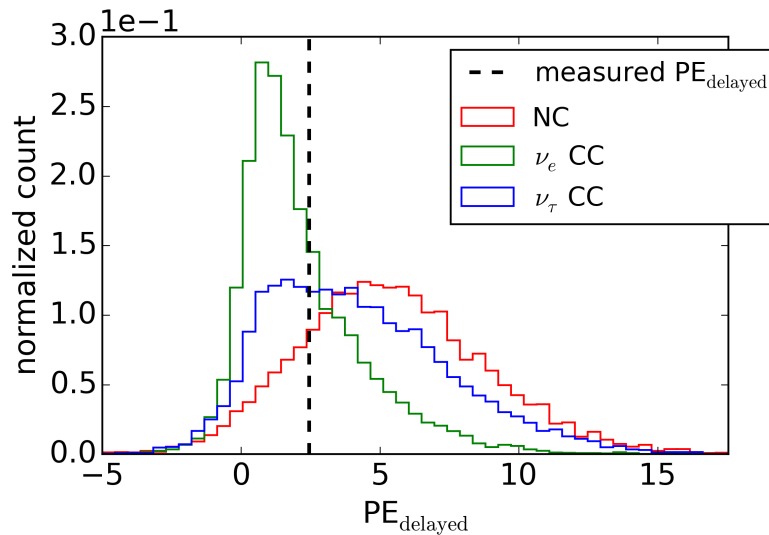

Figure E.7: $\text{PE}_{\text{delayed}}$ distributions for the event hypothesis of the HESE shower recorded on August 6, 2017. The dashed line marks the best-fit $\text{PE}_{\text{delayed}}$.

138

	iCDF($PE_{delayed,fit}$)
ν_e CC	0.75
ν_τ CC	0.94
NC	1.00

Table E.6: Value of the inverse $PE_{delayed}$-CDFs at the best-fit $PE_{delayed}$ for the HESE shower recorded on August 6, 2017.

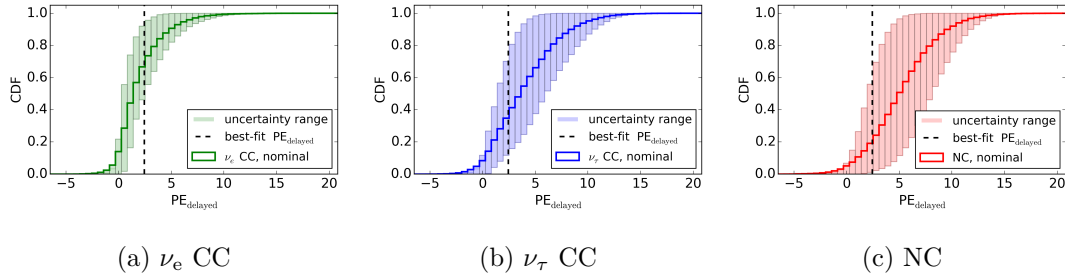(a) ν_e CC   (b) ν_τ CC   (c) NC

Figure E.8: CDFs corresponding to the distributions in Figure E.7. The bands show the $\pm 56\,\%$ systematic uncertainty range.

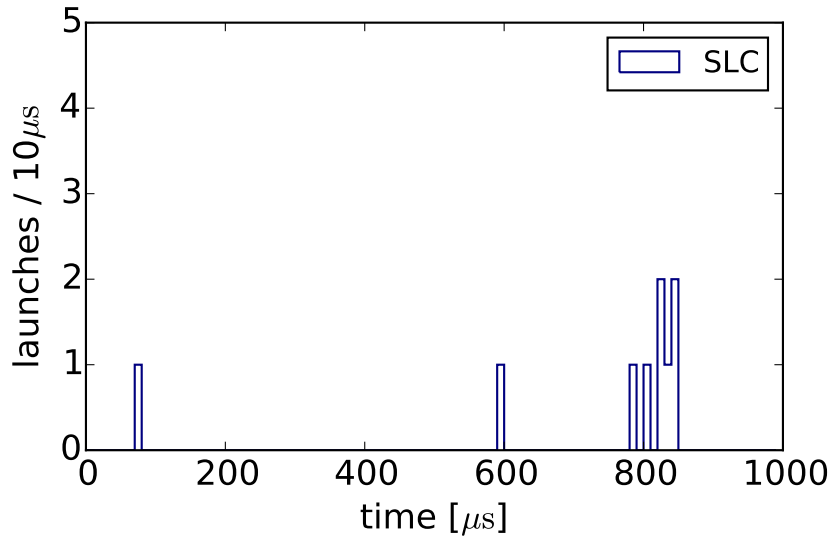Figure E.9: Delayed launch spectrum of the HESE shower recorded on August 6, 2017. 5.24 noise launches are expected in the $30 - 1000\,\mu s$ time window.

Event 4

event date	best-fit $PE_{delayed}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2017/06/26	46	397 TeV	29 m	-295 m

Table E.7: Delayed light deposition and event information of the HESE shower recorded on June 26, 2017. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
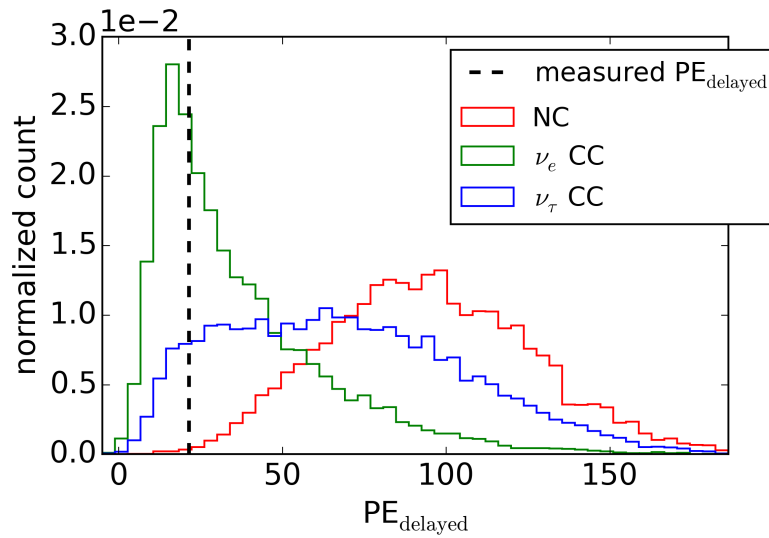

Figure E.10: $PE_{delayed}$ distributions for the event hypothesis of the HESE shower recorded on June 26, 2017. The dashed line marks the best-fit $PE_{delayed}$.

	iCDF(PE$_{\text{delayed,fit}}$)
ν_{e} CC	0.13
ν_{τ} CC	0.51
NC	0.84

Table E.8: Value of the inverse PE$_{\text{delayed}}$-CDFs at the best-fit PE$_{\text{delayed}}$ for the HESE shower recorded on June 26, 2017.

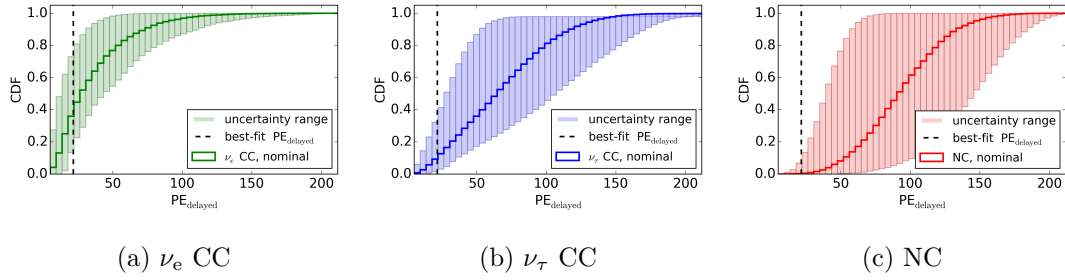(a) ν_{e} CC (b) ν_{τ} CC (c) NC

Figure E.11: CDFs corresponding to the distributions in Figure E.10. The bands show the $\pm 56\,\%$ systematic uncertainty range.
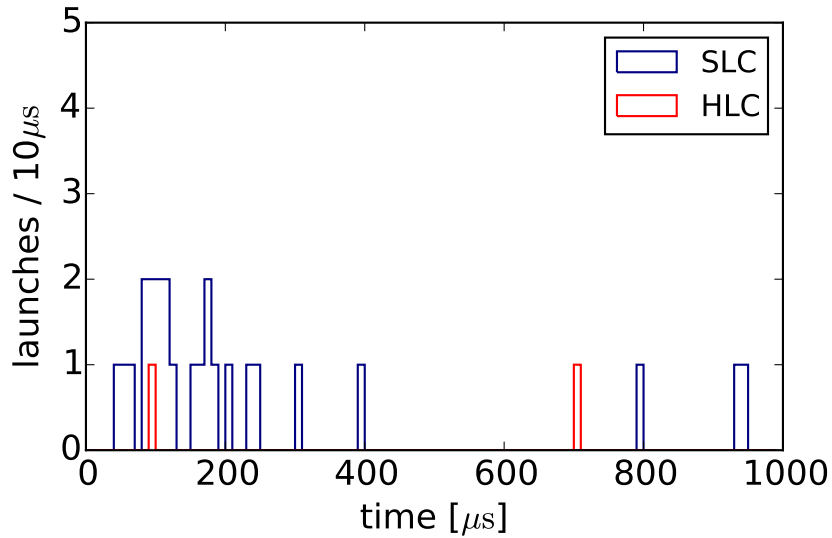

Figure E.12: Delayed launch spectrum of the HESE shower recorded on June 26, 2017. 5.58 noise launches are expected in the $30 - 1000\,\mu$s time window.

Event 5

event date	best-fit $PE_{delayed}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2017/05/11	11	118 TeV	23 m	-340 m

Table E.9: Delayed light deposition and event information of the HESE shower recorded on May 11, 2017. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
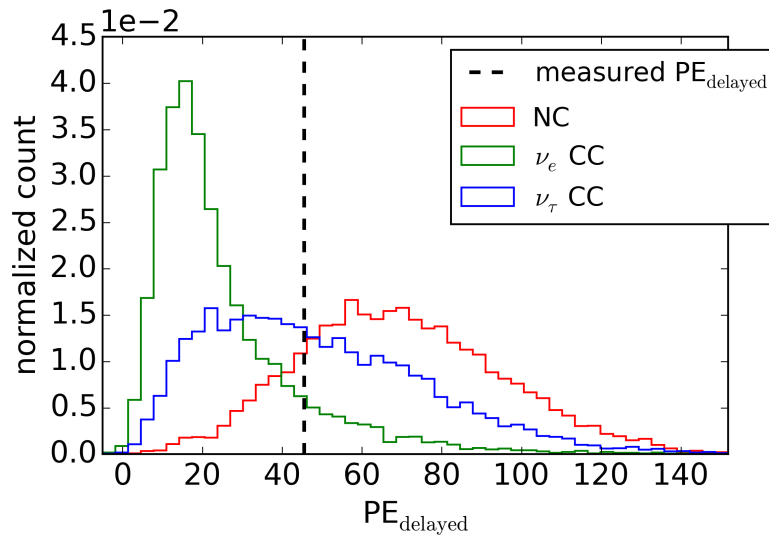

Figure E.13: $PE_{delayed}$ distributions for the event hypothesis of the HESE shower recorded on May 11, 2017. The dashed line marks the best-fit $PE_{delayed}$.

	iCDF($PE_{\text{delayed,fit}}$)
ν_e CC	0.23
ν_τ CC	0.59
NC	0.80

Table E.10: Value of the inverse PE_{delayed}-CDFs at the best-fit PE_{delayed} for the HESE shower recorded on May 11, 2017.

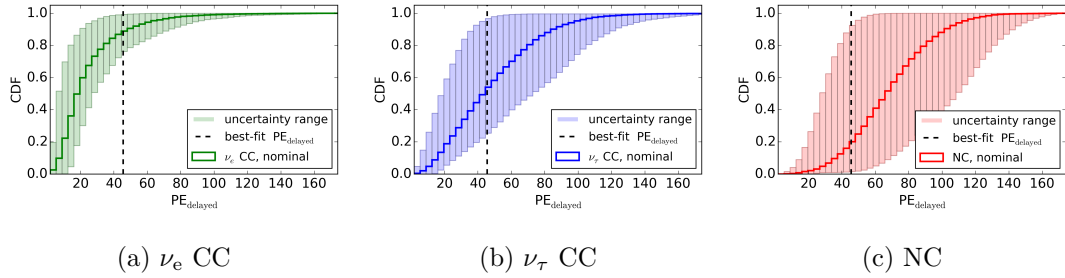(a) ν_e CC   (b) ν_τ CC   (c) NC

Figure E.14: CDFs corresponding to the distributions in Figure E.13. The bands show the $\pm 56\,\%$ systematic uncertainty range.
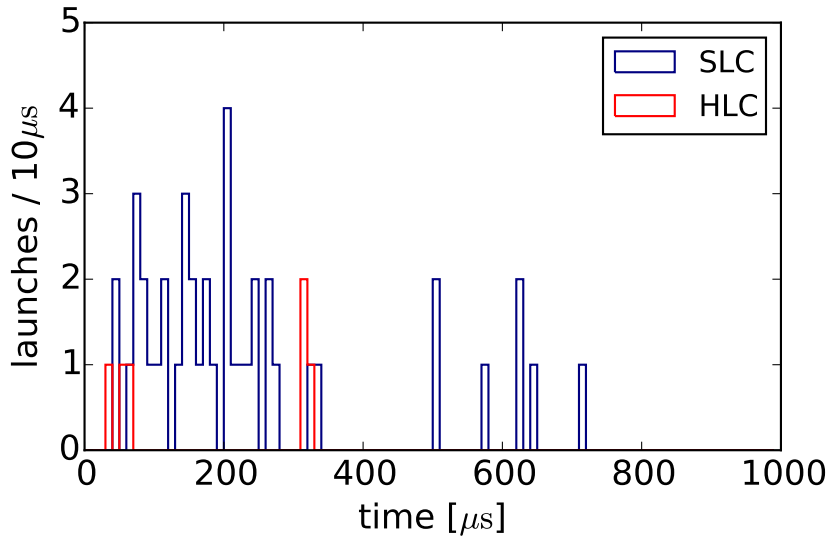

Figure E.15: Delayed launch spectrum of the HESE shower recorded on May 11, 2017. 3.71 noise launches are expected in the $30 - 1000\,\mu s$ time window.

Event 6

event date	best-fit $\mathrm{PE_{delayed}}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2017/03/23	1	82 TeV	49 m	-314 m

Table E.11: Delayed light deposition and event information of the HESE shower recorded on March 23, 2017. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
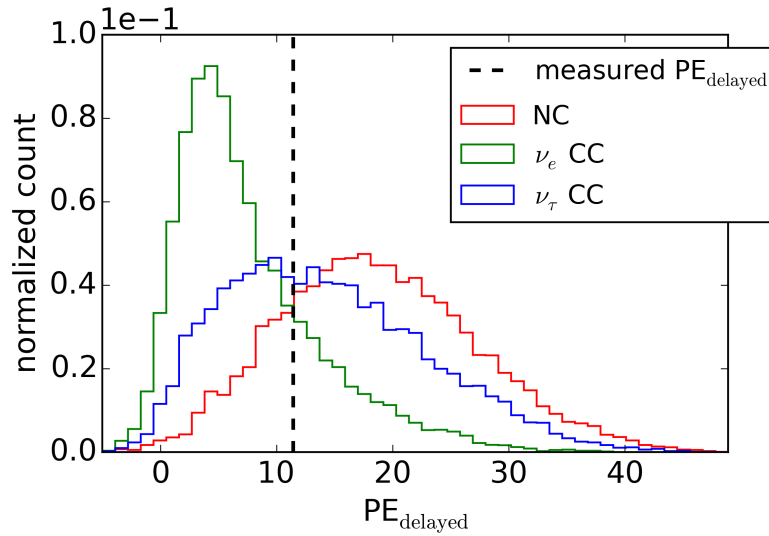

Figure E.16: $\mathrm{PE_{delayed}}$ distributions for the event hypothesis of the HESE shower recorded on March 23, 2017. The dashed line marks the best-fit $\mathrm{PE_{delayed}}$.

	iCDF($PE_{delayed,fit}$)
ν_e CC	0.77
ν_τ CC	0.87
NC	0.91

Table E.12: Value of the inverse $PE_{delayed}$-CDFs at the best-fit $PE_{delayed}$ for the HESE shower recorded on March 23, 2017.

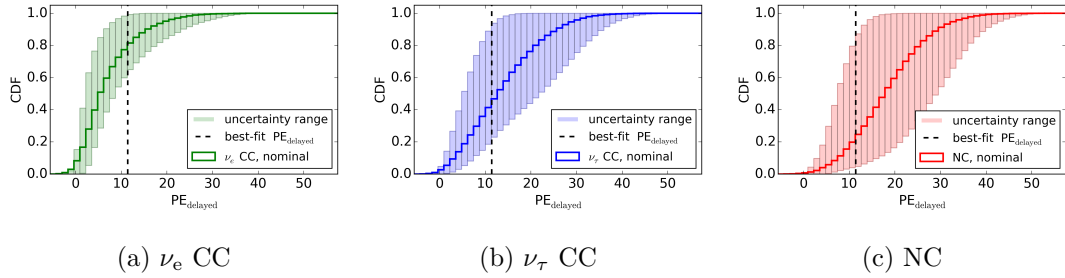(a) ν_e CC   (b) ν_τ CC   (c) NC

Figure E.17: CDFs corresponding to the distributions in Figure E.16. The bands show the $\pm 56\,\%$ systematic uncertainty range.
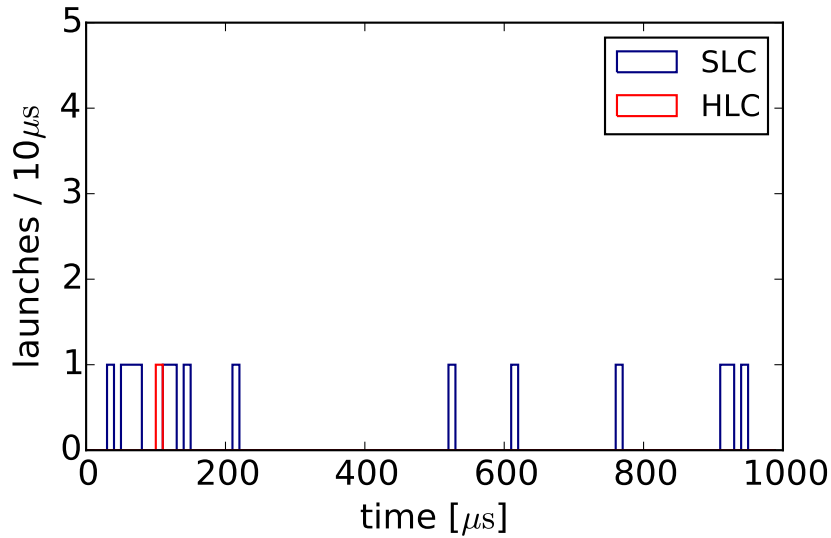

Figure E.18: Delayed launch spectrum of the HESE shower recorded on March 23, 2017. 3.76 noise launches are expected in the $30 - 1000\,\mu s$ time window.

Event 7

event date	best-fit $PE_{delayed}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2017/03/04	13	202 TeV	42 m	-446 m

Table E.13: Delayed light deposition and event information of the HESE shower recorded on March 4, 2017. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
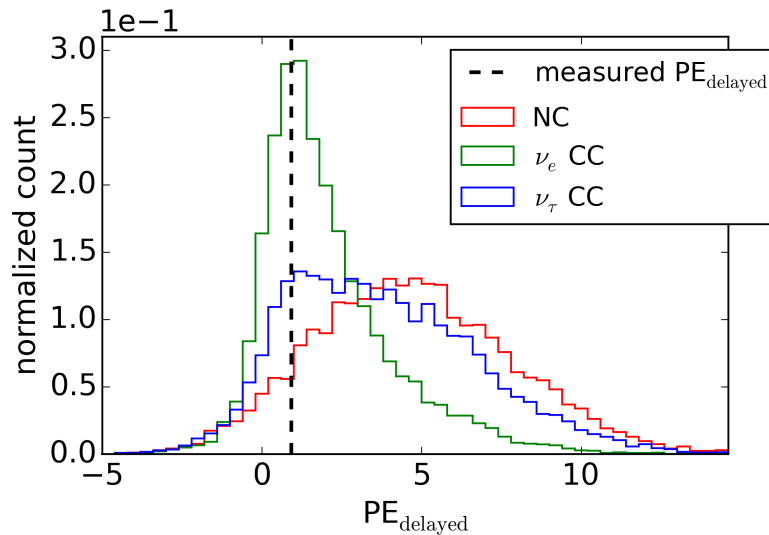

Figure E.19: $PE_{delayed}$ distributions for the event hypothesis of the HESE shower recorded on March 4, 2017. The dashed line marks the best-fit $PE_{delayed}$.

	iCDF($PE_{delayed,fit}$)
ν_e CC	0.04
ν_τ CC	0.22
NC	0.34

Table E.14: Value of the inverse $PE_{delayed}$-CDFs at the best-fit $PE_{delayed}$ for the HESE shower recorded on March 4, 2017.

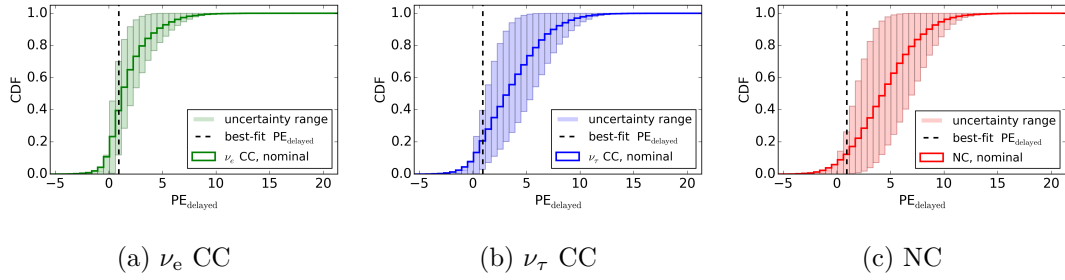(a) ν_e CC   (b) ν_τ CC   (c) NC

Figure E.20: CDFs corresponding to the distributions in Figure E.19. The bands show the $\pm 56\,\%$ systematic uncertainty range.
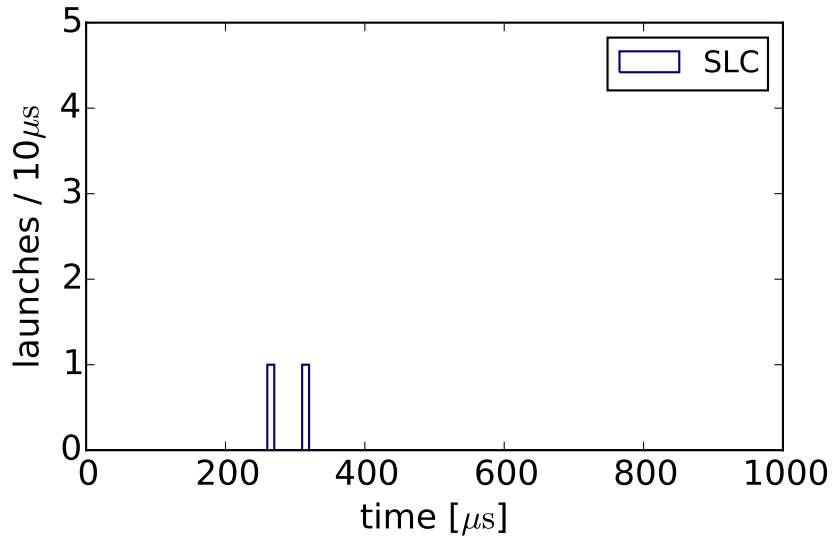

Figure E.21: Delayed launch spectrum of the HESE shower recorded on March 4, 2017. 3.61 noise launches are expected in the $30 - 1000\,\mu s$ time window.

Event 8

event date	best-fit $\mathrm{PE_{delayed}}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2016/12/26	4	165 TeV	45 m	185 m

Table E.15: Delayed light deposition and event information of the HESE shower recorded on December 26, 2016. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
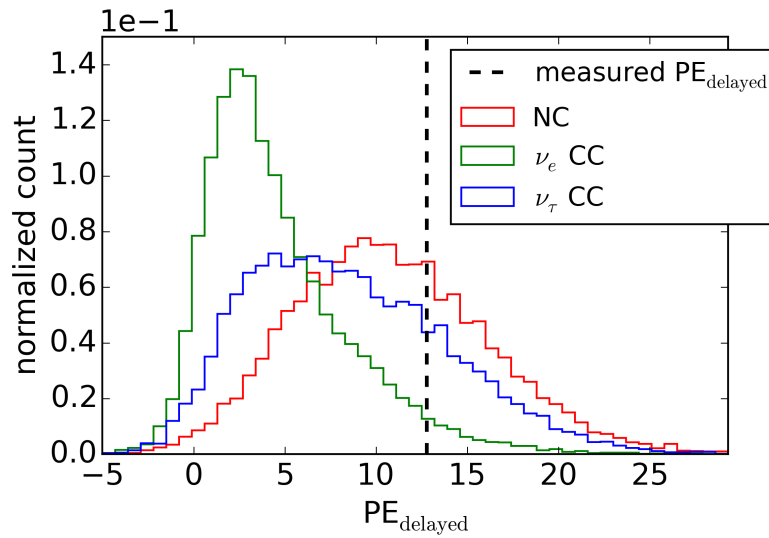

Figure E.22: $\mathrm{PE_{delayed}}$ distributions for the event hypothesis of the HESE shower recorded on December 26, 2016. The dashed line marks the best-fit $\mathrm{PE_{delayed}}$.

	iCDF(PE$_{\text{delayed,fit}}$)
ν_{e} CC	0.27
ν_{τ} CC	0.59
NC	0.78

Table E.16: Value of the inverse PE$_{\text{delayed}}$-CDFs at the best-fit PE$_{\text{delayed}}$ for the HESE shower recorded on December 26, 2016.

(a) ν_{e} CC   (b) ν_{τ} CC   (c) NC
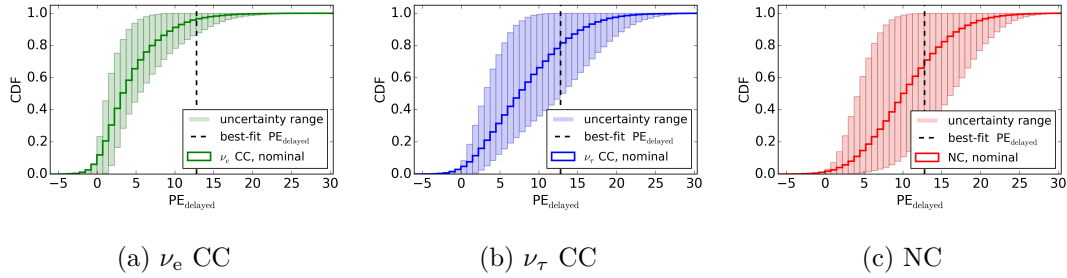
Figure E.23: CDFs corresponding to the distributions in Figure E.22. The bands show the $\pm 56\,\%$ systematic uncertainty range.
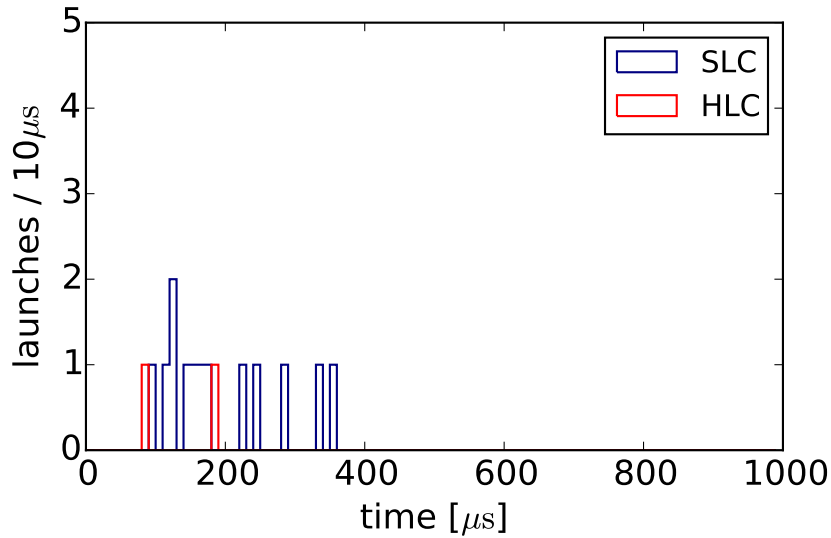

Figure E.24: Delayed launch spectrum of the HESE shower recorded on December 26, 2016. 4.64 noise launches are expected in the $30 - 1000\,\mu$s time window.

Event 9

event date	best-fit $PE_{delayed}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2016/11/30	10	260 TeV	38 m	107 m

Table E.17: Delayed light deposition and event information of the HESE shower recorded on November 30, 2016. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
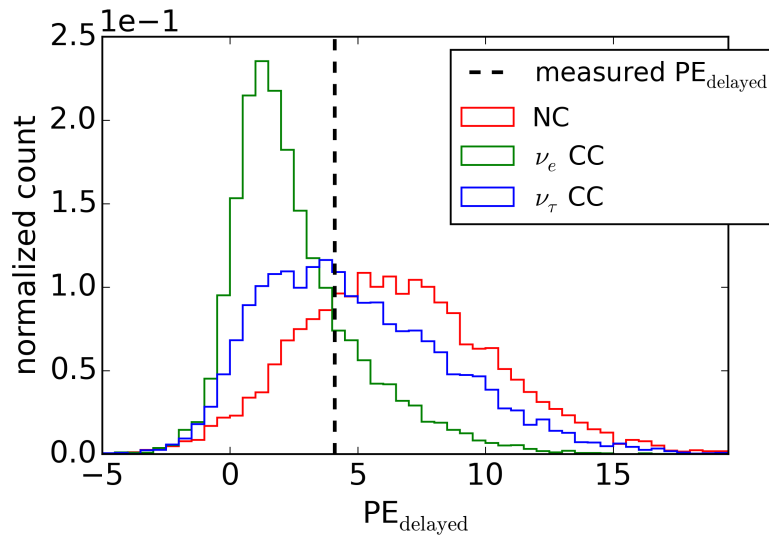

Figure E.25: $PE_{delayed}$ distributions for the event hypothesis of the HESE shower recorded on November 30, 2016. The dashed line marks the best-fit $PE_{delayed}$.

	iCDF(PE$_{\text{delayed,fit}}$)
ν_{e} CC	0.17
ν_{τ} CC	0.52
NC	0.78

Table E.18: Value of the inverse PE$_{\text{delayed}}$-CDFs at the best-fit PE$_{\text{delayed}}$ for the HESE shower recorded on November 30, 2016.

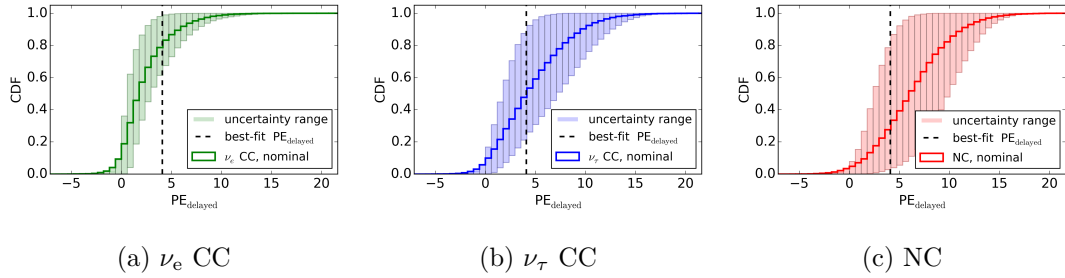(a) ν_{e} CC   (b) ν_{τ} CC   (c) NC

Figure E.26: CDFs corresponding to the distributions in Figure E.25. The bands show the $\pm 56\,\%$ systematic uncertainty range.
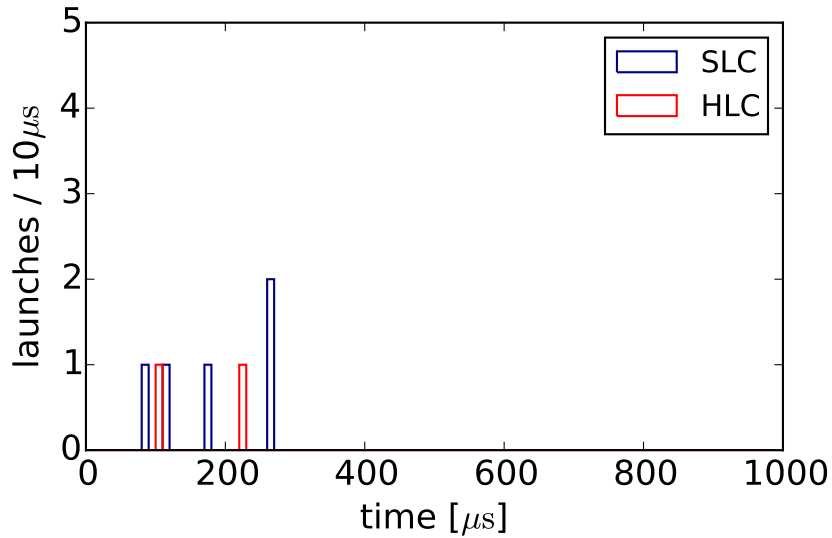

Figure E.27: Delayed launch spectrum of the HESE shower recorded on November 30, 2016. 4.39 noise launches are expected in the $30 - 1000\,\mu$s time window.

Event 10

event date	best-fit $PE_{delayed}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2016/11/06	4	65 TeV	39 m	50 m

Table E.19: Delayed light deposition and event information of the HESE shower recorded on November 6, 2016. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
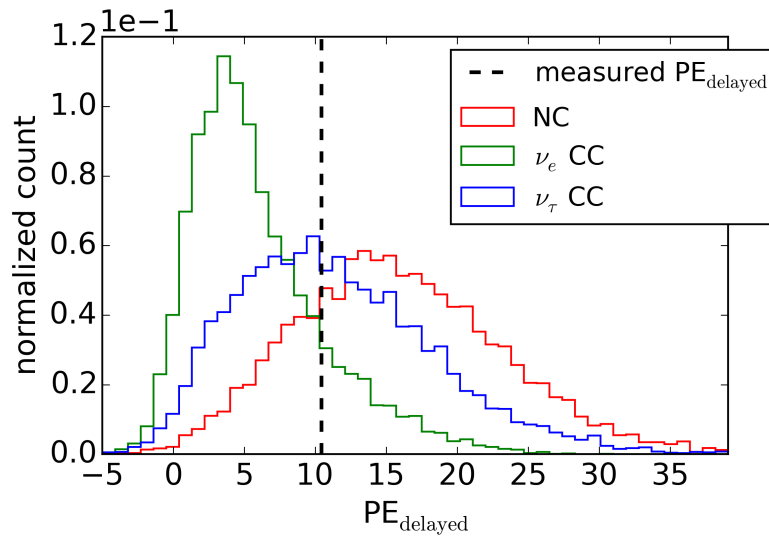

Figure E.28: $PE_{delayed}$ distributions for the event hypothesis of the HESE shower recorded on November 6, 2016. The dashed line marks the best-fit $PE_{delayed}$.

	iCDF(PE$_\text{delayed,fit}$)
ν_e CC	0.20
ν_τ CC	0.50
NC	0.70

Table E.20: Value of the inverse PE$_\text{delayed}$-CDFs at the best-fit PE$_\text{delayed}$ for the HESE shower recorded on November 6, 2016.

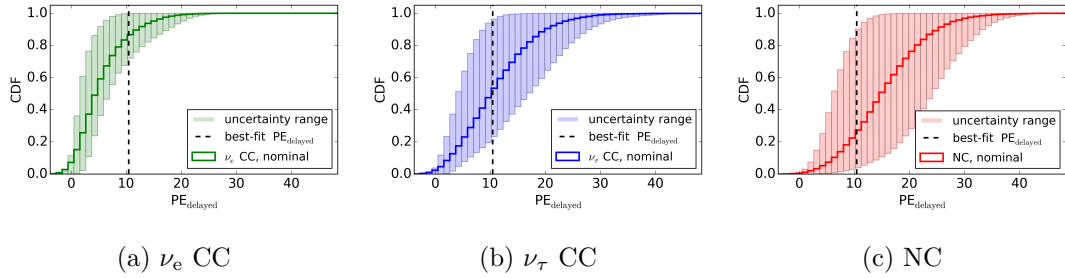(a) ν_e CC   (b) ν_τ CC   (c) NC

Figure E.29: CDFs corresponding to the distributions in Figure E.28. The bands show the $\pm 56\,\%$ systematic uncertainty range.
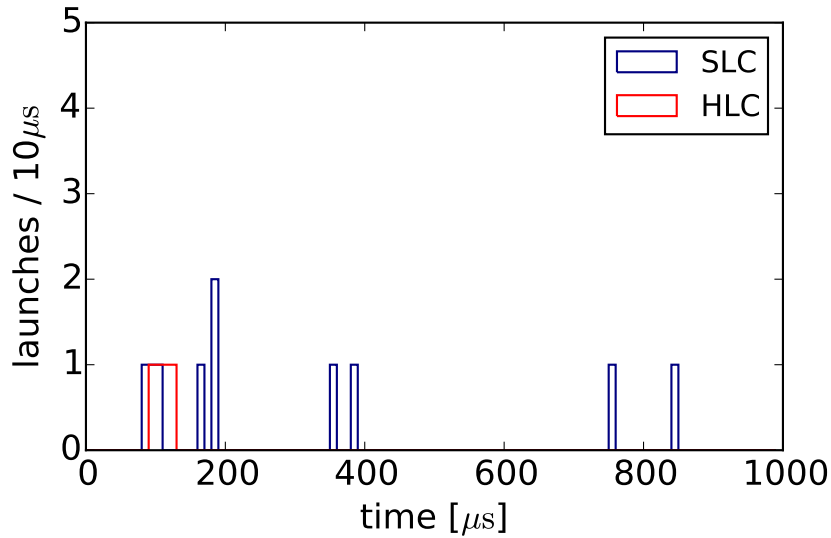

Figure E.30: Delayed launch spectrum of the HESE shower recorded on November 6, 2016. 4.02 noise launches are expected in the $30 - 1000\,\mu$s time window.

Event 11

event date	best-fit $\mathrm{PE_{delayed}}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2016/10/16	8	131 TeV	29 m	86 m

Table E.21: Delayed light deposition and event information of the HESE shower recorded on October 16, 2016. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
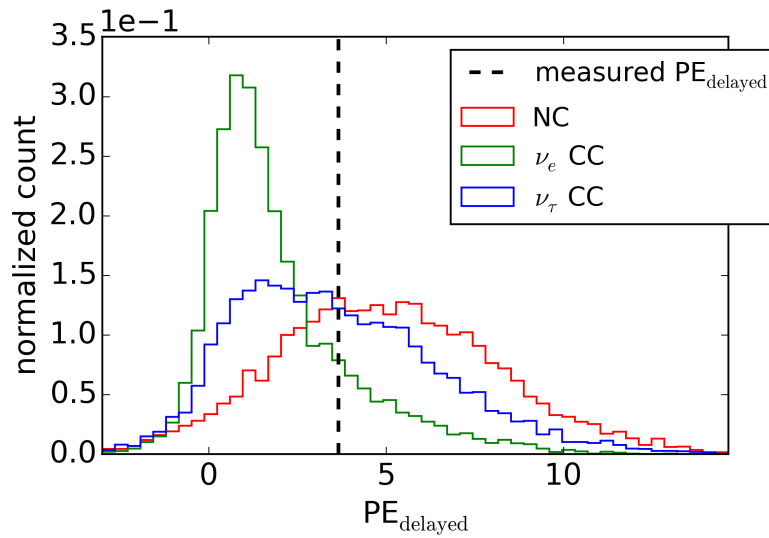

Figure E.31: $\mathrm{PE_{delayed}}$ distributions for the event hypothesis of the HESE shower recorded on October 16, 2016. The dashed line marks the best-fit $\mathrm{PE_{delayed}}$.

	iCDF(PE$_{\text{delayed,fit}}$)
ν_{e} CC	0.26
ν_{τ} CC	0.55
NC	0.81

Table E.22: Value of the inverse PE$_{\text{delayed}}$-CDFs at the best-fit PE$_{\text{delayed}}$ for the HESE shower recorded on October 16, 2016.

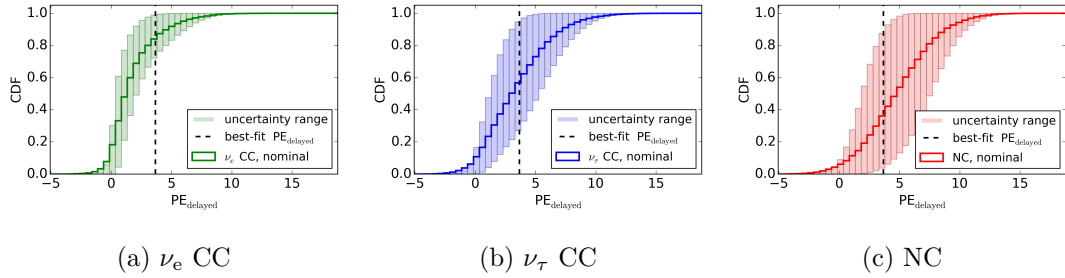(a) ν_{e} CC   (b) ν_{τ} CC   (c) NC

Figure E.32: CDFs corresponding to the distributions in Figure E.31. The bands show the $\pm 56\,\%$ systematic uncertainty range.
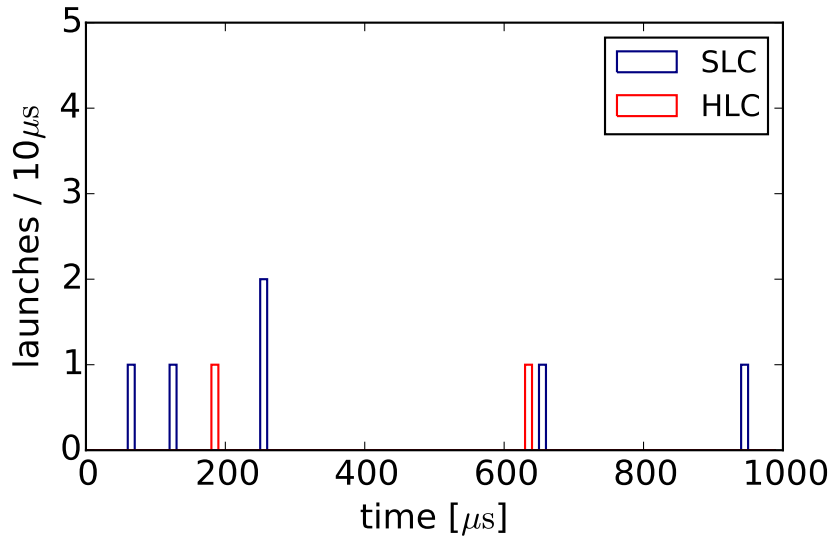

Figure E.33: Delayed launch spectrum of the HESE shower recorded on October 16, 2016. 4.18 noise launches are expected in the $30 - 1000\,\mu$s time window.

155

Event 12

event date	best-fit $PE_{delayed}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2016/10/06	22	132 TeV	38 m	-393 m

Table E.23: Delayed light deposition and event information of the HESE shower recorded on October 6, 2016. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
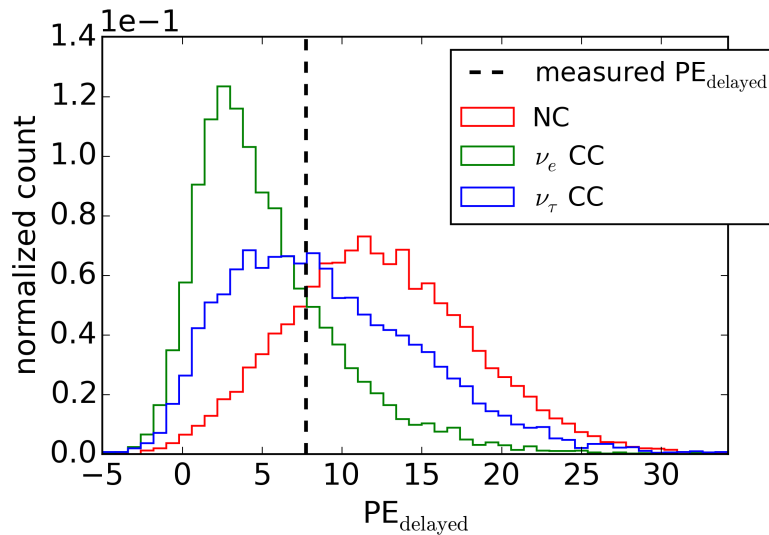

Figure E.34: $PE_{delayed}$ distributions for the event hypothesis of the HESE shower recorded on October 6, 2016. The dashed line marks the best-fit $PE_{delayed}$.

	iCDF(PE$_{\text{delayed,fit}}$)
ν_{e} CC	0.00
ν_{τ} CC	0.02
NC	0.03

Table E.24: Value of the inverse PE$_{\text{delayed}}$-CDFs at the best-fit PE$_{\text{delayed}}$ for the HESE shower recorded on October 6, 2016.

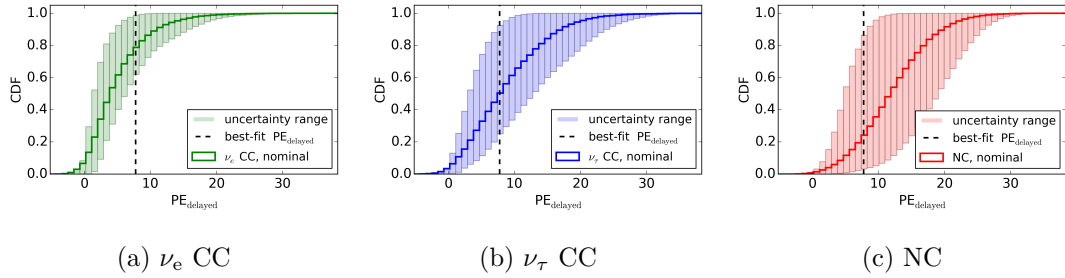(a) ν_{e} CC

(b) ν_{τ} CC

(c) NC

Figure E.35: CDFs corresponding to the distributions in Figure E.34. The bands show the $\pm 56\,\%$ systematic uncertainty range.
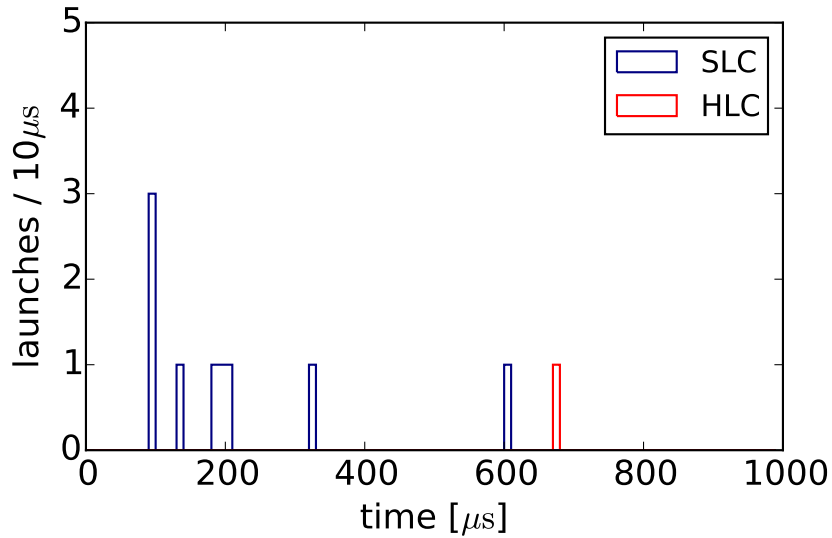

Figure E.36: Delayed launch spectrum of the HESE shower recorded on October 6, 2016. 4.29 noise launches are expected in the $30 - 1000\,\mu$s time window.

Event 13

event date	best-fit $PE_{delayed}$	reconstructed energy	distance between vertex and closest DOM	reconstructed vertex-z
2016/07/11	12	622 TeV	56 m	242 m

Table E.25: Delayed light deposition and event information of the HESE shower recorded on July 11, 2016. *Reconstructed vertex-z* refers to the z-coordinate of the reconstructed vertex position in IceCube coordinates, for which the center of the detector defines the origin of the coordinate system.
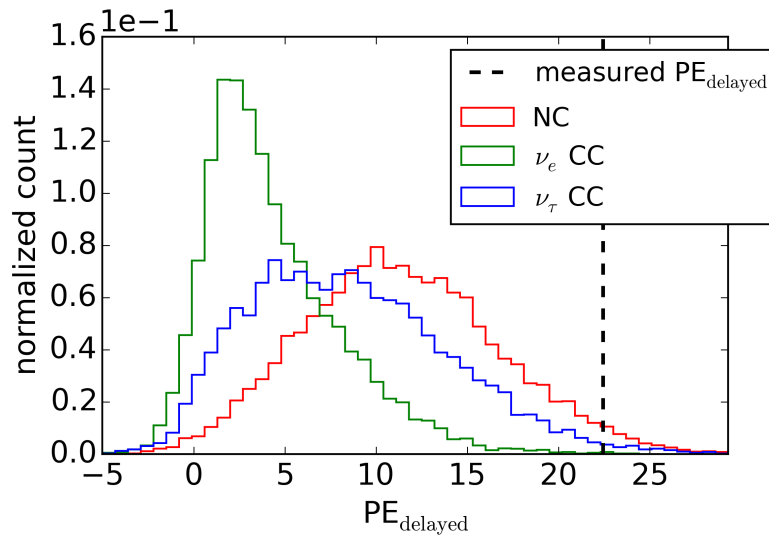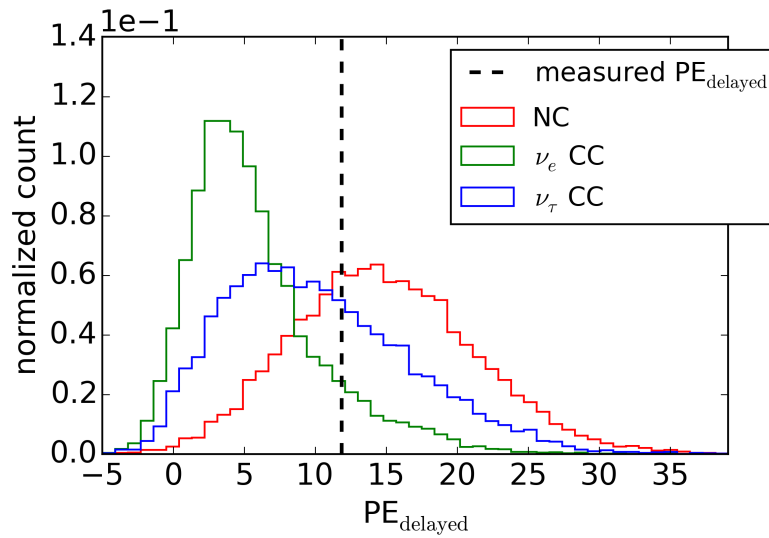

Figure E.37: $PE_{delayed}$ distributions for the event hypothesis of the HESE shower recorded on July 11, 2016. The dashed line marks the best-fit $PE_{delayed}$.

	$\mathrm{iCDF(PE_{delayed,fit})}$
ν_e CC	0.12
ν_τ CC	0.39
NC	0.69

Table E.26: Value of the inverse $\mathrm{PE_{delayed}}$-CDFs at the best-fit $\mathrm{PE_{delayed}}$ for the HESE shower recorded on July 11, 2016.

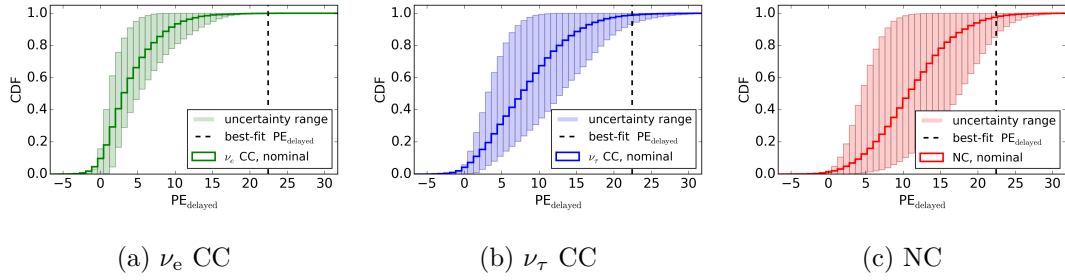(a) ν_e CC   (b) ν_τ CC   (c) NC
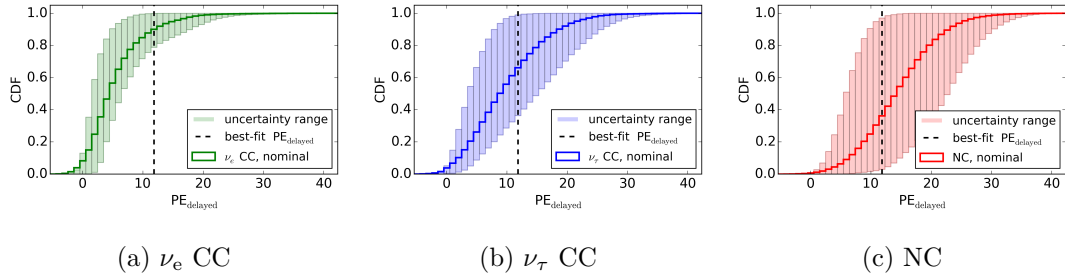
Figure E.38: CDFs corresponding to the distributions in Figure E.37. The bands show the $\pm 56\,\%$ systematic uncertainty range.
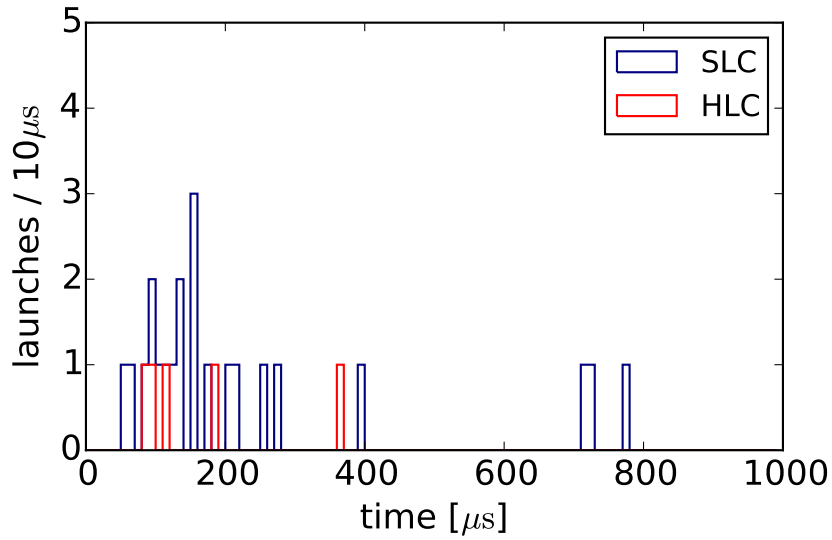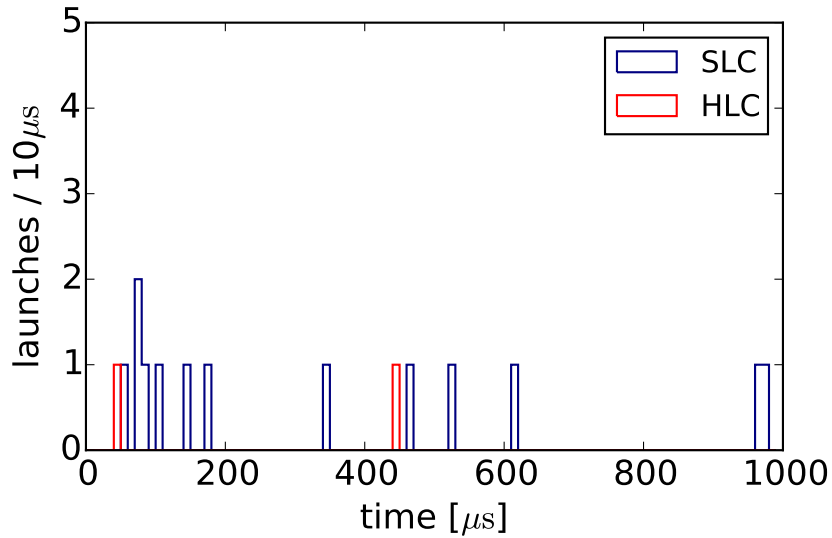

Figure E.39: Delayed launch spectrum of the HESE shower recorded on July 11, 2016. 4.32 noise launches are expected in the $30-1000\,\mu\mathrm{s}$ time window.

Bibliography

[1] J. Lesgourgues, S. Pastor, *Neutrino cosmology and Planck*, New J. Phys. 16, 065002 (2014).

[2] Y. Fukuda, et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. 81, 1562 (1998).

[3] B. Kayser, *Neutrino physics as explored by flavor change*, Phys. Rev. D66(010001), 392 (2002).

[4] F. Capozzi, et al., *Global constraints on absolute neutrino masses and their ordering*, Phys. Rev. D95(9), 096014 (2017).

[5] S. F. King, *Unified Models of Neutrinos, Flavour and CP Violation*, Prog. Part. Nucl. Phys. 94, 217 (2017).

[6] J. K. Becker, *High-energy neutrinos in the context of multimessenger physics*, Phys. Rept. 458, 173 (2008).

[7] C. Giunti, C. W. Kim, *Coherence of neutrino oscillations in the wave packet approach*, Phys. Rev. D58, 017301 (1998).

[8] V. Baum, *Search for Low Energetic Neutrino Signals from Galactic Supernovae and Collisionally Heated Gamma-Ray Bursts with the IceCube Neutrino Observatory*, Ph.D. thesis, Johannes Gutenberg-Universität Mainz (2017).

[9] R. Gandhi, et al., *Neutrino interactions at ultrahigh energies*, Phys. Rev. D 58, 093009 (1998).

[10] J. A. Formaggio, G. P. Zeller, *From eV to EeV: Neutrino Cross Sections Across Energy Scales*, Rev. Mod. Phys. 84, 1307 (2012).

[11] S. L. Glashow, *Resonant Scattering of Antineutrinos*, Phys. Rev. 118, 316 (1960).

[12] A. V. Butkevich, et al., *Ultrahigh-energy Neutrino Nucleon and Neutrino Electron Cross-sections in the Standard Model, in Supersymmetric and Superstring Models*, Z. Phys. C39, 241 (1988).

[13] R. Gandhi, et al., *Ultrahigh-energy neutrino interactions*, Astropart. Phys. 5, 81 (1996).

[14] L. Rädel, C. Wiebusch, *Calculation of the Cherenkov light yield from low energetic secondary particles accompanying high-energy muons in ice and water with Geant4 simulations*, Astropart. Phys. 38, 53 (2012).

160

[15] I. Tamm, *Radiation Emitted by Uniformly Moving Electrons*, pp. 37–53. Springer Berlin Heidelberg, Berlin, Heidelberg (1991).

[16] C. Patrignani, *Review of Particle Physics*, Chin. Phys. C40(10), 100001 (2016).

[17] E. Segrè, *Nuclei and Particles*. W.A. Benjamin (1964).

[18] *Atomic and Nuclear Properties of Materials for more than 300 materials, Particle Data Group Website*, `http://pdg.lbl.gov/2015/AtomicNuclearProperties/`, accessed April 13, 2016.

[19] L. Gerhardt, S. R. Klein, *Electron and Photon Interactions in the Regime of Strong LPM Suppression*, Phys. Rev. D82, 074017 (2010).

[20] L. Rädel, *Simulation Studies of the Cherenkov Light Yield from Relativistic Particles in High-Energy Neutrino Telescopes with Geant4*, Master's thesis, Rheinisch-Westfälische Technische Hochschule Aachen (2012).

[21] J. van Santen, *Neutrino Interactions in IceCube above 1 TeV: Constraints on Atmospheric Charmed-Meson Production and Investigation of the Astrophysical Neutrino Flux with 2 Years of IceCube Data taken 2010-2012*, Ph.D. thesis, University of Wisconsin-Madison (2014).

[22] M. Kowalski, *On the Cherenkov light emission of hadronic and electromagnetic cascades*, `http://internal.icecube.wisc.edu/reports/amanda/data/20020803-track.pdf` (2002), Internal Report 20020803, AMANDA Collaboration.

[23] M. G. Aartsen, et al., *Search for Astrophysical Tau Neutrinos in Three Years of IceCube Data*, Phys. Rev. D93(2), 022001 (2016).

[24] M. Usner, *Search for Astrophysical Tau Neutrinos in Six Years of High-Energy Starting Events in IceCube*, PoS ICRC2017, 974 (2017).

[25] S. W. Li, M. Bustamante, J. F. Beacom, *Echo Technique to Distinguish Flavors of Astrophysical Neutrinos* (2016).

[26] J. Bressieux, *Testing and comparison of muon energy estimators for the IceCube neutrino observatory*, Master's thesis, University of Wisconsin-Madison (2009).

[27] S. I. Klimushin, E. V. Bugaev, I. A. Sokalski, *Precise parametrizations of muon energy losses in water*, in *27th International Cosmic Ray Conference (ICRC 2001) Hamburg, Germany, August 7-15, 2001*, p. 1009 (2001), [3,1009(2001)].

[28] D. Bailey, *Monte Carlo tools and analysis methods for understanding the ANTARES experiment and predicting its sensitivity to dark matter*, Ph.D. thesis, Oxford U. (2002).

[29] M. G. Aartsen, et al., *Characterization of the Atmospheric Muon Flux in IceCube*, Astropart. Phys. 78, 1 (2016).

[30] R. Abbasi, et al., *The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping*, Nucl. Instrum. Meth. A 601, 294 (2009).

[31] R. Abbasi, et al., *The Design and Performance of IceCube DeepCore*, Astropart. Phys. 35, 615 (2012).

[32] M. G. Aartsen, et al., *The IceCube Neutrino Observatory: Instrumentation and Online Systems*, JINST 12(03), P03012 (2017).

[33] M. Wallraff, *Design, Implementation and Test of a New Feature Extractor for the IceCube Neutrino Observatory*, Diploma thesis, Rheinisch-Westfälische Technische Hochschule Aachen (2010).

[34] M. G. Aartsen, et al., *Energy Reconstruction Methods in the IceCube Neutrino Telescope*, JINST 9, P03009 (2014).

[35] R. Abbasi, et al., *Calibration and Characterization of the IceCube Photomultiplier Tube*, Nucl. Instrum. Meth. A618, 139 (2010).

[36] K. J. Ma, et al., *Time and Amplitude of Afterpulse Measured with a Large Size Photomultiplier Tube*, Nucl. Instrum. Meth. A629, 93 (2011).

[37] R. Abbasi, et al., *IceCube Sensitivity for Low-Energy Neutrinos from Nearby Supernovae*, Astron. Astrophys. 535, A109 (2011).

[38] D. Freiherr Heereman von Zuydtwyck, *HitSpooling: An Improvement for the Supernova Neutrino Detection System in IceCube*, Ph.D. thesis, Université Libre de Bruxelles (2015).

[39] S. R. Klein, *IceCube: A Cubic Kilometer Radiation Detector*, IEEE Trans. Nucl. Sci. 56, 1141 (2009).

[40] M. G. Aartsen, et al., *Search for a diffuse flux of astrophysical muon neutrinos with the IceCube 59-string configuration*, Phys. Rev. D89(6), 062007 (2014).

[41] M. G. Aartsen, et al., *Search for neutrino-induced particle showers with IceCube-40*, Phys. Rev. D89(10), 102001 (2014).

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

[43] M. G. Aartsen, et al., *Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector*, Science 342, 1242856 (2013).

[44] M. G. Aartsen, et al., *Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data*, Phys. Rev. Lett. 113, 101101 (2014).

162

[45] C. Kopper, W. Giang, N. Kurahashi, *Observation of Astrophysical Neutrinos in Four Years of IceCube Data*, PoS ICRC2015, 1081 (2016).

[46] *Observation of Astrophysical Neutrinos in Six Years of IceCube Data*, PoS ICRC2017, 981 (2017).

[47] S. Choubey, W. Rodejohann, *Flavor Composition of UHE Neutrinos at Source and at Neutrino Telescopes*, Phys. Rev. D80, 113006 (2009).

[48] M. G. Aartsen, et al., *A combined maximum-likelihood analysis of the high-energy astrophysical neutrino flux measured with IceCube*, Astrophys. J. 809(1), 98 (2015).

[49] E. Fermi, *On the Origin of the Cosmic Radiation*, Phys. Rev. 75, 1169 (1949).

[50] A. R. Bell, *The Acceleration of cosmic rays in shock fronts. I*, Monthly Notices of the Royal Astronomical Society 182(2), 147 (1978).

[51] T. K. Gaisser, T. Stanev, *High-energy cosmic rays*, Nucl. Phys. A777, 98 (2006).

[52] H. Niederhausen, Y. Xu, *High Energy Astrophysical Neutrino Flux Measurement Using Neutrino-induced Cascades Observed in 4 Years of IceCube Data*, PoS ICRC2017, 968 (2017).

[53] C. Haack, C. Wiebusch, *A Measurement of the Diffuse Astrophysical Muon Neutrino Flux Using Eight Years of IceCube Data*, PoS ICRC2017, 1005 (2017).

[54] L. Lu, *A Measurement of the Diffuse Astrophysical Muon Neutrino Flux Using Eight Years of IceCube Data*, PoS ICRC2017, 1002 (2017).

[55] J. Söngen, *Nachweis von verzögerten Photonen in IceCube*, Bachelor's thesis, Johannes Gutenberg-Universität Mainz (2015).

[56] Neutron News 3(3), 29 (1992).

[57] S. Hirschberg, *Neutron slowing-down time in finite water systems*, Annals of Nuclear Energy (Oxford) 10(8), 405 (1983).

[58] E. Morris, *A theoretical analysis of the neutron scattering method of measuring snow and ice density* 113 (2008).

[59] *AntarcticGlaciers.org*, http://www.antarcticglaciers.org/glaciers-and-climate/ice-cores/ice-core-basics/, accessed April 19, 2017.

[60] J. Voß, *Konzept einer Neutronen-Kalibrations-Sonde zum Einsatz im tiefen antarktischen Eis*, Bachelor's thesis, Johannes Gutenberg-Universität Mainz (2017).

[61] L. Köpke, *Neutrons from hadronic interactions in ice* (2017), Internal Report icecube/201707001-v3, IceCube Collaboration.

[62] A. Pollmann, S. Pieper, *Measurement of water luminescence - a new detection method for neutrino telescopes*, PoS ICRC2017, 1060 (2017).

[63] T. I. Q. S. M. Trotman, D. F. Sangster, *Decay kinetics of the ultraviolet and visible luminescences emitted by electron-irradiated crystalline H_2O ice*, The Journal of Chemical Physics 85, 2555 (1986).

[64] T. Quickenden, T. Green, D. Lennon, *Luminescence from UV-Irradiated Amorphous H_2O Ice*, The Journal of Physical Chemistry 42, 16801 (1996).

[65] A. Pollmann, private correspondence (2017).

[66] S. W. Li, J. F. Beacom, *First calculation of cosmic-ray muon spallation backgrounds for MeV astrophysical neutrino signals in Super-Kamiokande*, Phys. Rev. C89, 045801 (2014).

[67] S. W. Li, J. F. Beacom, *Spallation Backgrounds in Super-Kamiokande Are Made in Muon-Induced Showers*, Phys. Rev. D91(10), 105005 (2015).

[68] J. Haser, et al., *Afterpulse Measurements of R7081 Photomultipliers for the Double Chooz Experiment*, JINST 8, P04029 (2013).

[69] M. Zoll, *Preparations for the next solar WIMP Analysis with IceCube*, Licentiate thesis, Stockholm University (2014).

[70] M. Zoll, *A search for solar dark matter with the IceCube neutrino detector*, Ph.D. thesis, Stockholm University (2016).

[71] M. Larson, *Simulation and identification of non-Poissonian noise triggers in the IceCube neutrino detector*, Master's thesis, University of Alabama (2013).

[72] J. Lundberg, et al., *Light tracking for glaciers and oceans: Scattering and absorption in heterogeneous media with Photonics*, Nucl. Instrum. Meth. A581, 619 (2007).

[73] N. Whitehorn, J. van Santen, S. Lafebre, *Penalized Splines for Smooth Representation of High-dimensional Monte Carlo Datasets*, Comput. Phys. Commun. 184, 2214 (2013).

[74] T. Sjöstrand, et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. 191, 159 (2015).

[75] T. Sjöstrand, personal communication (2016).

[76] S. Höche, S. Prestel, *The midpoint between dipole and parton showers*, The European Physical Journal C 75(9), 461 (2015).

[77] M. Dobbs, J. B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, Comput. Phys. Commun. 134, 41 (2001).

[78] S. Agostinelli, et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506(3), 250 (2003).

[79] J. Allison, et al., *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. 53, 270 (2006).

[80] J. Allison, et al., *Recent developments in Geant4*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 835, 186 (2016).

[81] P. Martinerie, et al., *Air content paleo record in the Vostok ice core (Antarctica): A mixed record of climatic and glaciological parameters*, Journal of Geophysical Research Atmospheres 99(D5), 10565 (1994).

[82] L. Grevillot, et al., *Optimization of GEANT4 settings for Proton Pencil Beam Scanning simulations using GATE*, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 268(20), 3295 (2010).

[83] *Physics Reference Manual, Version: geant4 10.1*, `http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/PhysicsReferenceManual/BackupVersions/V10.1/fo/PhysicsReferenceManual.pdf` (2014).

[84] *clsim*, `https://github.com/claudiok/clsim`, accessed April 25, 2017.

[85] D. Chirkin, *Evidence of optical anisotropy of the South Pole ice*, in *Proceedings, 33rd International Cosmic Ray Conference (ICRC2013): Rio de Janeiro, Brazil, July 2-9, 2013*, p. 0580.

[86] A. C. Meinusch, *Nachweis verzögerter Neutronen in IceCube*, Bachelor's thesis, Johannes Gutenberg-Universität Mainz (2017).

[87] A. C. Meinusch, L. Köpke, *IceCube deadtime effects on long lasting signals* (2017), Internal Report icecube/201706001, IceCube Collaboration.

[88] D. Pandel, *Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal Neutrinoteleskop NT-72*, Diploma thesis, Humboldt Universität zu Berlin (1996).

[89] G. Bellini, et al., *Cosmogenic Backgrounds in Borexino at 3800 m water-equivalent depth*, JCAP 1308, 049 (2013).

[90] S. Abe, et al., *Production of Radioactive Isotopes through Cosmic Muon Spallation in KamLAND*, Phys. Rev. C81, 025807 (2010).

[91] Y. Zhang, et al., *First measurement of radioactive isotope production through cosmic-ray muon spallation in Super-Kamiokande IV*, Phys. Rev. D93(1), 012004 (2016).

[92] Y. Chen, S. Ahmad, *SU-GG-T-333: Neutron Yield and Angular Distribution From the 250 MeV Proton Interactions in Water: A Geant4 Monte Carlo Study*, Medical Physics 35(6), 2802 (2008).

[93] A. Howard, et al., *Validation of neutrons in Geant4 using TARC data - production, interaction and transportation*, in *Proceedings, 2008 IEEE Nuclear Science Symposium, Medical Imaging Conference and 16th International Workshop on Room-Temperature Semiconductor X-Ray and Gamma-Ray Detectors (NSS/MIC 2008 / RTSD 2008): Dresden, Germany, October 19-25, 2008*, pp. 2885–2889 (2008).

[94] *Minuit*, `http://seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/`, accessed February 16, 2018.

[95] H. Niederhausen, *Performance study of Cascade Reconstruction at the Glashow Resonance in IceCube*, Master's thesis, Stony Brook University (2012).

[96] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. 9(1), 60 (1938).

[97] T. Bayes, *LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S*, Philosophical Transactions 53, 370 (1763).

[98] D. Foreman-Mackey, et al., *emcee: The MCMC Hammer*, Publications of the ASP 125, 306 (2013).

[99] N. Metropolis, et al., *Equation of State Calculations by Fast Computing Machines*, The Journal of Chemical Physics 21(6), 1087 (1953).

[100] W. K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika 57(1), 97 (1970).

[101] D. Foreman-Mackey, et al., *corner.py: corner.py v2.0.0* (2016).

[102] F. James, *Statistical Methods in Experimental Physics*. World Scientific, 2nd edn. (2006).

[103] J. Kopp, J. Liu, X.-P. Wang, *Boosted Dark Matter in IceCube and at the Galactic Center*, JHEP 04, 105 (2015).

[104] A. Bhattacharya, R. Gandhi, A. Gupta, *The Direct Detection of Boosted Dark Matter at High Energies and PeV events at IceCube*, JCAP 1503(03), 027 (2015).

[105] A. Bhattacharya, et al., *Boosted Dark Matter and its implications for the features in IceCube HESE data*, JCAP 1705(05), 002 (2017).

[106] L. Necib, et al., *Boosted Dark Matter at Neutrino Experiments*, Phys. Rev. D95(7), 075018 (2017).

[107] C. Kachulis, et al., *Search for Boosted Dark Matter Interacting With Electrons in Super-Kamiokande* (2017).

[108] H. Zhang, et al., *Supernova Relic Neutrino Search with Neutron Tagging at Super-Kamiokande-IV*, Astropart. Phys. 60, 41 (2015).

[109] Y. S. Jeong, et al., *Tau energy loss and ultrahigh energy skimming tau neutrinos*, Phys. Rev. D96(4), 043003 (2017).

[110] I. Anghel, et al., *Letter of Intent: The Accelerator Neutrino Neutron Interaction Experiment (ANNIE)* (2015).

[111] D. Schmitt, *Nature of astrophysical neutrinos: shower identification by the delayed neutron echo*, Bachelor's thesis, Johannes Gutenberg-Universität Mainz (2018).

[112] M. G. Aartsen, et al., *IceCube-Gen2: A Vision for the Future of Neutrino Astronomy in Antarctica*, PoS FRAPWS2016, 004 (2017).

[113] *GEANT4 User Support Pages*, `http://geant4.cern.ch/support/index.shtml`, accessed April 12, 2017.