Measurement of the High Energy Astrophysical Neutrino Flux Using Electron and Tau Neutrinos Observed in Four Years of IceCube Data

A Dissertation Presented

by

Hans Niederhausen

to

The Graduate School

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Physics

Stony Brook University

May 2018

Stony Brook University

The Graduate School

Hans Niederhausen

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

Joanna Kiryluk - Dissertation Advisor
Associate Professor, Department of Physics and Astronomy

Dominik A. Schneble - Chairperson of Defense
Associate Professor, Department of Physics and Astronomy

Dmitri Kharzeev
Professor, Department of Physics and Astronomy

David Kawall
Associate Professor, Department of Physics
University of Massachusetts Amherst

This dissertation is accepted by the Graduate School.

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Measurement of the High Energy Astrophysical Neutrino Flux Using Electron and Tau Neutrinos Observed in Four Years of IceCube Data

by

Hans Niederhausen

Doctor of Philosophy

in

Physics

Stony Brook University

2018

The high-energy universe is known to be violent. Ultra High Energy Cosmic Rays (UHECRs) have been observed with kinetic energies exceeding 10^{20} eV. Their origin, despite decades of observations, remains elusive. A unique probe of the sources and production mechanisms of these high energy cosmic rays can be neutrinos, since they are inevitably produced when high-energy protons interact. The IceCube Neutrino Observatory, located at the geographical South Pole in Antarctica, continuously monitors a total volume of $1\,\mathrm{km}^3$ of clear Antarctic ice for neutrino interactions. For this purpose, a total of 5160 optical sensors (photomultiplier tubes) have been melted deep into the glacier at depths between 1450 m and 2450 m. In 2013 IceCube reported one of its biggest discoveries, the observation of highly energetic neutrinos that are consistent with a possible extra-galactic origin.

In this dissertation we use IceCube data (recorded from 2012 to 2015) to study the spectral properties of this astrophysical neutrino flux with focus on electron and tau neutrino flavors. We developed a new neutrino identification and muon background rejection method using state-of-the-art machine-learning techniques, more specifically multi-class gradient boosted decision trees. In addition to enlarging the number of detected neutrino events ($>$ 10x increase over previous works), we lowered the energy threshold to below 1 TeV and thereby greatly improved upon the control and treatment of systematic uncertainties. The sample contains $\sim$ 400 astrophysical electron and tau neutrinos, which increases the significance of the original discovery to 8.7σ. We find the astrophysical neutrino flux to be

well described by a single power-law consistent with expectations from Fermi-type acceleration of high-energy particles at astrophysical sources and obtain leading constraints on its properties. In particular this dataset favors a spectral index $\gamma = 2.53^{+0.07}_{-0.09}$ and a flux normalization per ν-flavor of $\Phi = (1.58^{+0.25}_{-0.28}) \cdot 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$ at $\mathrm{E}_\nu = 100\,\mathrm{TeV}$. We further studied the possibility of additional spectral complexity, which significantly increases measurement uncertainties. No evidence for such scenarios was found. Finally we searched for a contribution from atmospheric neutrinos related to heavy meson (charm) decay in Earth's atmosphere and derive a flux upper limit of 4.8 times the benchmark pQCD flux prediction at 90% confidence level, dominated by systematic uncertainties, especially related to photon transport in the glacial ice.

Acknowledgements

First and foremost I would like to express my sincerest gratitude to my advisor Dr. Joanna Kiryluk for her invaluable guidance and counsel throughout the years. Without her most helpful advice this work would not have reached the finish line. I very much appreciate the fruitful collaboration with Yiqian Xu at Stony Brook University, especially during the development of this event selection and study of IceCube simulation data. I would like to further thank my dissertation committee: Dr. Dominik Schneble, Dr. Dmitri Kharzeev and Dr. David Kawall for their time and patience reviewing this manuscript. Similarly I appreciate the time spent by Dr. Dawn Williams and Dr. Gisela Anton throughly inspecting this analysis as part of IceCube's internal review. I am also thankful for Dr. Nancy Wandkowsky's efforts concerning simulation production in IceCube, which this work has greatly benefited from. Finally, I would like to thank the entire IceCube Collaboration for their support.
Schlussendlich möchte ich mich vielmals bei meinen Eltern für die fortwährende Unterstützung bedanken. Ohne euch hätte ich diese Arbeit nicht zustande gebracht.

The author of this dissertation would like to thank Stony Brook Research Computing and Cyberinfrastructure, and the Institute for Advanced Computational Science at Stony Brook University for access to the SeaWulf computing system, which was made possible by a $1.4M National Science Foundation grant (#1531492).

The material presented in this dissertation is based upon work supported by the National Science Foundation under Grant No. 1205796 and 1555121.

Contents

List of Figures

List of Tables

Chapter 1

Particle Astrophysics and High Energy Neutrinos

Cosmic radiation that penetrates Earth's atmosphere from outer space has been observed since the early 20th century [1]. Now, more than 100 years later, different types of experiments routinely measure its physical properties. By now, the cosmic radiation has been shown to consist of charged particles and nuclei, called cosmic rays, which were discovered first. Subsequently, the existence of cosmic γ-rays, a neutral component, was established in the 1960s [2][3]. Around the same time the first neutrinos of extra-terrestrial origin were found: neutrino emission from the sun [4]. The first detection of neutrinos from beyond our solar system, emitted by supernova explosion SN 1987A, was achieved in 1987. In 2012 the IceCube experiment discovered a diffuse flux of high energy, astrophysical neutrinos [5]. Each of these three[1] cosmic messengers adds unique information that can help to solve (some) of the existing puzzles in particle astrophysics. Despite continuous scientific progress throughout the last century, the pressing question: "What are the sources and production mechanisms of high energy cosmic particles" remains a topic of on-going research.

1.1 High Energy Cosmic Rays

Cosmic rays consist of ionized nuclei with a negligible contribution from electrons $< 1\%$ [8], which we will thus ignore in the context of this work. Roughly 90% of the nuclei are protons, 9% are helium nuclei (α-particles) and the rest heavier nuclei. The differential flux of cosmic rays dN/dE decreases with the kinetic energy of the nucleus (primary particle[2]) and, to a good approximation, follows a doubly broken power law $dN/dE \propto E^{-\gamma}$ over ~ 9 orders of magnitude. The spectral indices are observed as

- $\gamma \approx 2.7$ for $10\,\mathrm{GeV} \lesssim E \lesssim 3 \cdot 10^6\,\mathrm{GeV}$

[1] A fourth cosmic messenger would be gravitational waves, discovered recently by LIGO [6].

[2] Throughout this work we will use the term particle and nucleus interchangeably. We will instead use the word elementary particle to refer to fundamental particles without further substructure.

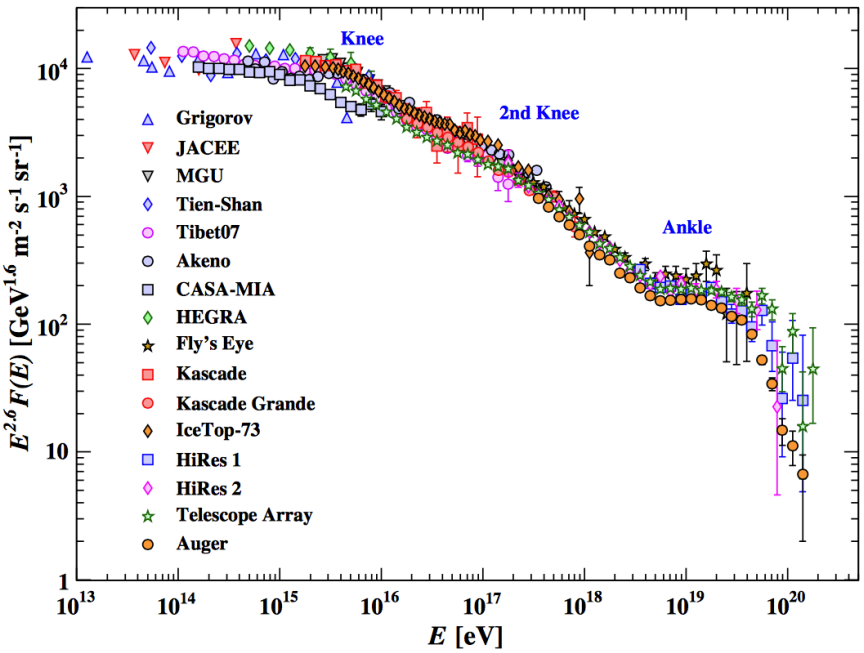


FIGURE 1.1: The flux of high energy cosmic-rays (all particles) as function of energy-per-nucleus obtained from air-shower measurements (Figure from [7]).

- $\gamma \approx 3.1$ for $3 \cdot 10^6\,\text{GeV} \lesssim E \lesssim 3 \cdot 10^9\,\text{GeV}$ ("knee")
- $\gamma \approx 2.6$ for $3 \cdot 10^9\,\text{GeV} \lesssim E$ ("ankle").

The flux appears to cut-off (vanish rapidly) above $\sim 4 \cdot 10^{10}\,\text{GeV}$ [9][10]. At low to moderate energies $< 100\,\text{TeV}$ the flux of cosmic rays (as a function of energy E and atomic number Z of the primary nucleus) can be measured directly using balloon (e.g. CREAM [11], TRACER[12]) and satellite-borne (e.g. PAMELA[13], AMS-02[14]) particle detectors. Since the flux decreases rapidly with energy, such direct experiments do not provide sufficient collection area to accumulate cosmic-ray events above 100 TeV and thus large area, ground based experiments are required. These air-shower experiments (e.g. Auger[9], Telescope Array[10], IceTop[15]) rely on the detection of secondary particles produced in extensive air-showers initiated by collisions between cosmic-ray nucleons and nucleons in Earth's atmosphere. The properties of the primary cosmic-rays have to be inferred from the observed properties of the particle showers on the ground. However, not being able to measure the primary cosmic-rays directly, severely limits the ability of such air-shower detectors to distinguish between different nuclei. Hence there are large uncertainties about the composition of the cosmic-ray flux at highest energies.
In the context of production of secondary particles in cosmic-ray air-showers it will be useful to parametrize the cosmic-ray flux in terms of the *energy-per-nucleon* E_N, since cosmic-ray interactions essentially reduce to nucleon-nucleon collisions in Earth's atmosphere. Below the cosmic-ray knee $E < 1\,\text{PeV}$ the differential nucleon flux is [7]

$$\frac{dN}{dE_N} = 1.8 \times 10^4 \left(\frac{E_N}{\text{GeV}}\right)^{-2.7} \frac{\text{nucleons}}{\text{m}^2\,\text{s}\,\text{sr}\,\text{GeV}} \tag{1.1}$$

Free protons contribute the majority of the nucleons to the flux ($\sim 74\%$) followed by helium nuclei ($\sim 18\%$).
The presence of knee ($\sim 3\,\mathrm{PeV} - 3\,\mathrm{EeV}$) and ankle ($\sim 3\,\mathrm{EeV} - 30\,\mathrm{EeV}$) in the energy spectrum as well the observed suppression of the flux at the high energy end ($E \sim 40\,\mathrm{EeV}$) indicate changes in the underlying population of cosmic rays and require explanation. In the absence of experimental identification of cosmic-ray sources at all energies, assumptions have to be made. Galactic supernovae are believed to be the dominant source of high energy cosmic rays below the knee. First, the expanding supernova shock has the potential to "sweep up" material from the interstellar medium and efficiently accelerate particles to high energies via *diffusive shock acceleration*. Second, the energy output $\sim 10^{42}\,\mathrm{erg/s}$ from galactic supernovae, contributing to the expanding shock, would be sufficient to counteract the loss of cosmic-rays and corresponding loss of energy from our galaxy if a supernova rate of $\sim 3 \cdot (100\,\mathrm{y})^{-1}$ is assumed [16] . In addition this requires that a fraction $10^{-2} < \eta < 10^{-1}$ of that energy could be used to accelerate particles. DSA can provide such conversion efficiency. Furthermore, the chemical composition of the observed cosmic rays is consistent with expectations from a supernova origin and stellar nucleosynthesis. Finally, recent γ-ray observations of *supernova remnants* (SNRs) support the idea that protons are present in such shocks [17].
While probably contributing significantly to the observed flux of cosmic rays, c.f. Fig. 1.1, SNRs can not be responsible for all of it. Using DSA (c.f Sec. 1.2) it is possible to derive an estimate of the *maximum energy* beyond which SNRs can no longer accelerate the particles. Assuming a galactic magnetic field of $\sim 3\mu G$ and a typical SNR lifespan of $\sim 10^3\,\mathrm{y}$ one finds [18][19]:

$$E_{max} \leq Z \times 3 \times 10^4\,\mathrm{GeV} \tag{1.2}$$

More recent calculations obtain somewhat higher estimates. SNRs could potentially accelerate protons up to $\sim 10^{15}\,\mathrm{GeV}$ [20][16] and thus reach the on-set of the knee in the cosmic-ray spectrum. This can be pushed further if one considers the possibility that magnetic fields in the vicinity of the shock might be enhanced compared to the the galactic average [19][16].
It is natural to assume that cosmic accelerators can be characterized by a maximum energy E_{max}. Charged particles with atomic number Z and energy E are confined by magnetic fields with strength B. The characteristic length scale for spiral motion is given by the *gyro-radius* r_g

$$r_g = \frac{R}{Bc} = 1.08 \frac{E/\mathrm{PeV}}{Z \cdot B/\mu\mathrm{G}}\,\mathrm{pc} \quad \text{with} \quad R \equiv \frac{E}{Ze} \tag{1.3}$$

where R is called rigidity and characterizes the relevant particle properties. A cosmic accelerator can not confine particles with gyro-radii larger than its own size, thereby defining a maximum particle energy

$$E_{max} = Z \times \frac{R_s/\mathrm{pc} \cdot B/\mu\mathrm{G}}{1.08}\,\mathrm{PeV} \tag{1.4}$$

where R_s denotes the size of the acceleration region. Eq. (1.4) is called *Hillas criterion* [21]. Both eqs., (1.2) and (1.4), depend on Z and predict heavier elements to achieve higher energies. This provides an intuitive interpretation of the knee in the cosmic-ray

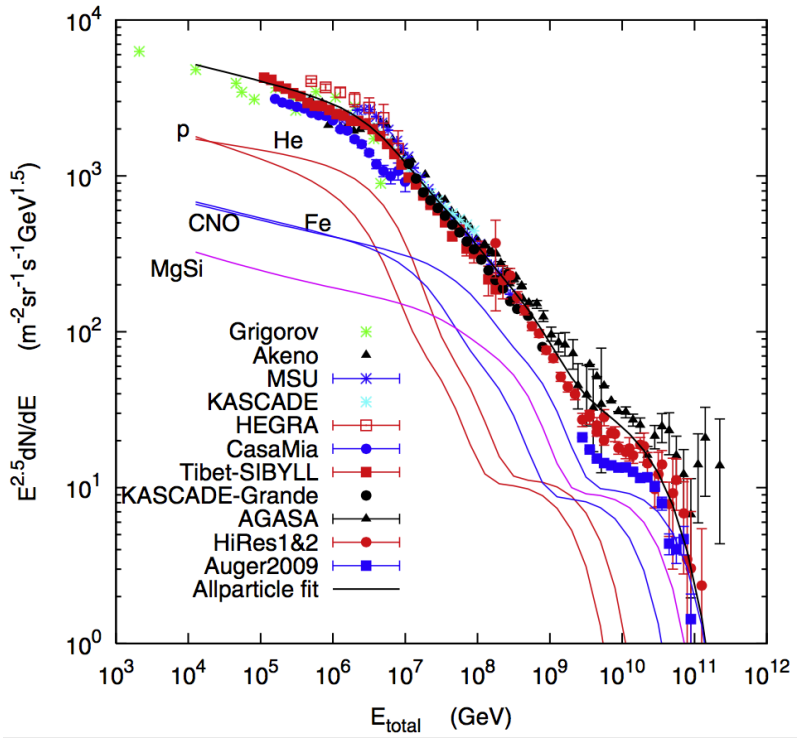


FIGURE 1.2: The Gaisser-H3a Cosmic Ray Flux Model [27] compared to air-shower measurements of the all-particle cosmic ray spectrum. (Figure from [27])

spectrum via the *Peters cycle* [22]. In this picture the on-set of the knee is given by the maximum energy for protons of the astrophysical accelerator. Heavier elements start to contribute more prominently as the flux is increasingly depleted from protons, until they too reach their maximum energy. Thus one expects heavy primaries (up to iron) to eventually dominate the flux. This is consistent with measurements by the KASCADE air-shower array [23]. Alternative scenarios, however, can also produce such a composition cycle, e.g. propagation effects like rigidity dependent leakage from the galaxy [19][24].

At highest energies $\gtrsim 10^{18}\,\mathrm{eV}$ the spectrum shows another spectral transition, the ankle. From eq. (1.4) one finds that such particles have gyro-radii $\gtrsim 300\,\mathrm{pc}$ of similar size than thickness of our galaxy. Hence they could not be contained within the galaxy. In addition, there are no galactic objects with a known mechanism that could explain such high energies. Thus cosmic-rays that contribute to the ankle are believed to be extra-galactic in origin [16]. Cosmic rays in this energy range are usually referred to as *ultra high energy cosmic rays* (UHECRs). The Auger [9] experiment recently reported the discovery of an anisotropy in the arrival directions of the UHECRs above 8 EeV in support of the extra-galactic interpretation [25].

Careful study of these two components, galactic SNRs and un-identified extra-galactic sources, have shown that a simple superposition of these two might not be sufficient to explain the high energy end of knee [26]. The proposed solution introduces a third component at intermediate energies: high energy cosmic rays produced in as of yet un-identified,

R_c	γ	p	He	CNO	Mg-Si	Fe
γ for Pop. 1	-	1.66	1.58	1.63	1.67	1.63
Population 1: 4 PV	see line 1	7860	3550	2200	1430	2120
Population 2: 30 PV	1.4	20	20	13.4	13.4	13.4
Pop. 3 (mixed): 2 EV	1.4	1.7	1.7	1.14	1.14	1.13
Pop. 3 (proton only): 60 EV	1.6	200.0	0	0	0	0

TABLE 1.1: Nominal Cosmic Ray Flux Model (Gaisser H3a/H4a): Cutoffs, integral spectral indices and normalizations constants $a_{i,j}$ for eq. (1.5). values from [27].

presumably galactic sources (potentially pulsars or micro-quasars [16]). However extra-galactic possibilities exist relaxing the need for an extra component [28][3]. This three component model has been fit to cosmic-ray data, combining direct detection methods (CREAM baloon), anchoring the low energy spectrum and composition, with air shower measurements that constrain the spectral behavior at highest energies [27]. To simplify the description of the composition, only the most abundant cosmic ray nuclei are described and, for this purpose, are grouped into five components: protons, helium, CNO, Mg-Si and Mn-Fe. The all particle spectrum then is described as the superposition of each component from each group - each parametrized as power-law with high-energy, rigidity-dependent exponential cutoff [27].

$$\Phi_E = \sum_{i=1}^{5} \sum_{j=1}^{3} a_{i,j} E^{-\gamma_{i,j}} \cdot \exp\left(-\frac{E}{Z_i \cdot R_{c,j}}\right) \tag{1.5}$$

The best-fit values for this model can be found in Tab. 1.1. Since the air-shower experiments only provide loose constraints on the composition at highest energies, two results are obtained: the *Gaisser H3a* parametrization assumes a mixed composition, while a pure proton composition is used in *Gaisser H4a* [30]. Fig. 1.2 compares the H3a flux for each mass group to the observed data. In this work we will use the H3a parametrization as baseline cosmic-ray flux model.

1.2 Diffusive Shock Acceleration

The detailed physics of particle acceleration in astrophysical sources remains an area of active research [31]. For reasons that we will outline below it is generally assumed that particle acceleration takes place in astrophysical plasma shocks. The argument goes back to Fermi, who showed that highly energetic particles can gain energy from elastic interactions with irregular magnetic fields embedded in moving clouds of plasma [32]. In particular, he derived that such processes lead to power law spectra for the accelerated particles in agreement with observations. The main problem with this *second order Fermi acceleration* involves efficiency considerations: typical velocities of interstellar clouds are small $\beta \equiv \frac{v}{c} \leq 10^{-4}$ and so is the probability for scattering of cosmic rays in the interstellar medium [33]. Thus, particles would gain energy too slowly. However his idea stimulated a lot of further theoretical developments that culminated in the late 1970s in a model called

[3]Such models exhibit potential conflict with non-observation of UHE neutrinos by IceCube [29].

first order Fermi acceleration or *diffusive shock acceleration* [34][35][36]. In this model particle acceleration is believed to occur at the front of hydrodynamical plasma shocks, as for example present in supernova remnants [37]. Here we will follow the exposition in [19]. For simplicity a planar, non-relativistic shock in an ionized gas (plasma) is considered with velocity parallel to the normal direction of the shock. In addition we require the presence of fully turbulent magnetic fields that will randomize the trajectories of highly relativistic charged particles.
The physics of shock waves is described by fluid dynamics [38]. A disturbance that propagates through a medium at a speed v_s greater than the speed of sound c_m in that medium will cause a shock with *Mach number* $M = v_s/c_m > 1$. A shock essentially is a solution to the fluid dynamics equations permitting a discontinuity in the flow of the fluid (here: plasma, i.e. hot ionized gas) [39]. The thermodynamical properties describing the flow of the medium behind the shock (*down-stream*) are influenced by the passage of the shock, while those of the flow ahead of it (*up-stream*) are not. The physical properties of shocks are best analyzed in its own rest-frame, i.e. a frame in which $v_s = 0$. The situation is visualized in Fig. 1.3 (top left). The up-stream (un-shocked) gas is moving into the shock front with relative speed u_1. Upon passage through the shock into the down-stream region, the gas departs from the shock front with reduced relative speed of $u_2 < u_1$. Requiring that mass be conserved at the shock and assuming an ideal gas one can calculate the ratio between both speeds, the *compression ratio* of the shock, as function of the Mach number [38]

$$\frac{u_1}{u_2} = \frac{r_c + 1}{r_c - 1 + 2/M^2} \tag{1.6}$$

where $r_c = c_p/c_v$ is the ratio between the specific heats of the gas at constant pressure c_p and constant volume c_v. $\gamma = 5/3$ for a monatomic, ideal gas [38].

Let us now assume that there exists a population of relativistic, charged particles near the front of the shock. These particles are very much faster than the shock itself and thus barely notice it. For simplicity we further assume the influence of this population on the shock itself to be negligible, i.e. we consider the *test particle regime*. Under these assumptions we will be able to show that a particle, after one complete cycle of crossing from the up-stream region of the shock into the down-stream region and subsequently crossing back into the up-stream region, on average gains energy. Given the much larger speed of the highly relativistic particle compared to that of the shock, such particle can undergo many of these cycles during the lifespan[4] of the shock and thus gain enormous amounts of energy from the plasma - provided it can be confined magnetically near the shock front for sufficient amount of time.
To make this quantitative consider a highly-relativistic particle ($E \sim pc$) propagating towards the shock from the up-stream side with energy E_1' and momentum projection onto the shock-normal direction $\boldsymbol{p_1}' \cdot \hat{\boldsymbol{n_s}} \equiv \boldsymbol{p_1}' \cdot \hat{\boldsymbol{x}} = E_1' \cos\theta_1'$. Here we use primed coordinates to denote measured values in a frame in which the up-stream gas is at rest. This plasma rest frame is identical with the lab frame and is shown in Fig. 1.3 (top right), with the trajectory of the test-particle marked green. The corresponding plasma velocities ($\mathbf{u_1}' = \mathbf{0}$, $\mathbf{u_2}' = (u_1 - u_2)\,\hat{\boldsymbol{x}}$) are related to the ones defined in the shock rest frame (un-primed) via Galilean transformation. An observer moving with the down-stream plasma

[4] characteristic timescales are $O(10^3)$ y for the shock front of a typical super nova remnant[19].

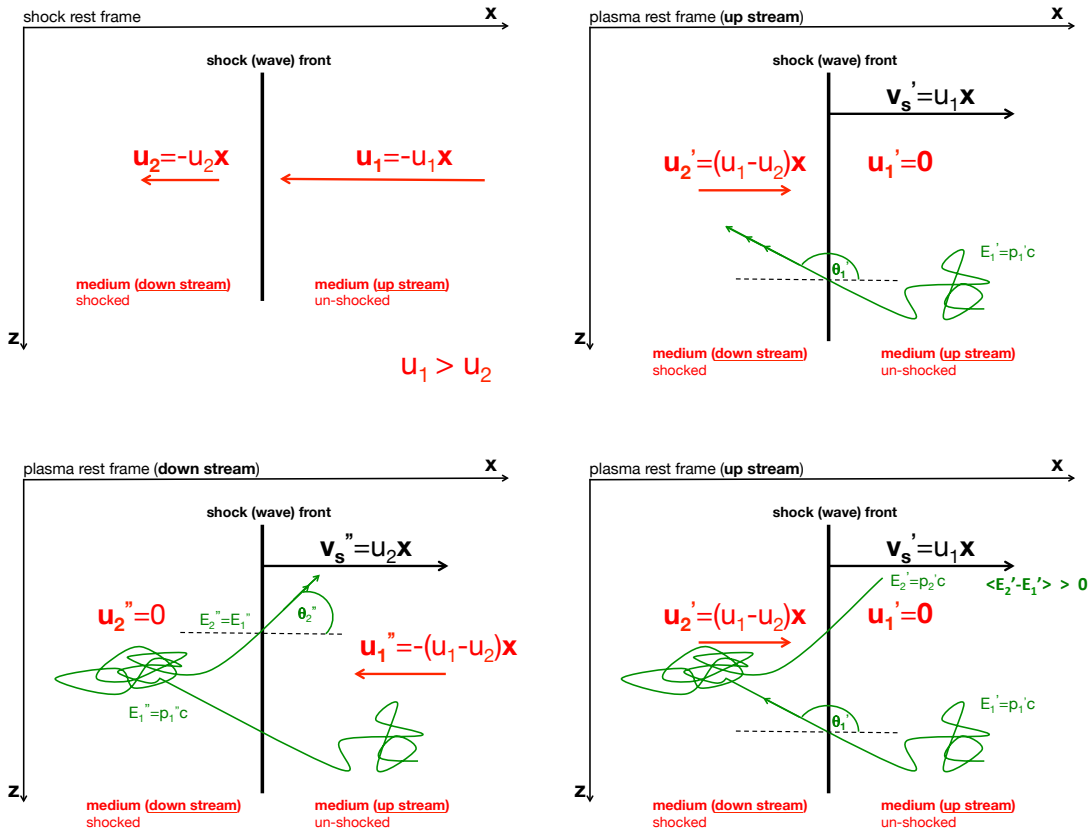


FIGURE 1.3: Diffusive Shock Acceleration described in different reference frames: shock rest frame (top left), rest frame of up-stream plasma (top right), rest frame of down stream plasma (bottom left) and rest frame of up-stream plasma (bottom right) after one complete cycle of particle crossings (see text). (Figure adapted from [16])

(double-primed coordinates) will measure a particle energy

$$E_1'' = \gamma E_1' \left(1 - \beta \cos\theta_1'\right) \tag{1.7}$$

where $\beta = (u_1 - u_2)/c$ denotes the relative velocity between both frames and thus the relative velocity between the plasmas on both sides of the shock, and $\gamma = (1-\beta^2)^{-\frac{1}{2}}$ is the corresponding Lorentz factor. This plasma rest-frame (down-stream) is shown in Fig. 1.3 (bottom left). Assuming that turbulent magnetic fields are present, the trajectory of the particle is quickly randomized. This is an elastic process and energy is neither lost nor gained. The particle crosses back into the up-stream region with energy $E_2'' = E_1''$ and therefore with momentum $\boldsymbol{p_2}'' \cdot \hat{\boldsymbol{x}} = E_1'' \cos\theta_2''$. An observer moving with the up-stream gas (Fig. 1.3 bottom right) sees the particle emerging from the shock with energy

$$E_2' = \gamma E_1'' \left(1 + \beta \cos\theta_2''\right) \tag{1.8}$$

Substituting E_1'' in eq. (1.8) by eq. (1.7) one finds

$$E_2' = \gamma^2 E_1' \left(1 + \beta \cos\theta_2''\right)\left(1 - \beta \cos\theta_1'\right) \tag{1.9}$$

Thus the relative energy gain for one cycle with angles θ_1' and θ_2'' becomes

$$\frac{\Delta E}{E} = \frac{E_2' - E_1'}{E_1'} = \frac{1 + \beta\cos\theta_2'' - \beta\cos\theta_1' - \beta^2\cos\theta_2''\cos\theta_1'}{1-\beta^2} - 1 \tag{1.10}$$

Any particular choice of angles is arbitrary and we are interested in the average gain ξ

$$\begin{aligned}\xi \equiv &<< \frac{\Delta E}{E} >_{\theta_1'} >_{\theta_2''} = \\ &\frac{1 + \beta < \cos\theta_2'' >_{\theta_2''} - \beta < \cos\theta_1' >_{\theta_1'} - \beta^2 < \cos\theta_1' >_{\theta_1'} < \cos\theta_2'' >_{\theta_2''}}{1-\beta^2} - 1\end{aligned} \tag{1.11}$$

To calculate these averages we need to abandon the single particle picture and consider an ensemble of particles of equal energies instead. We wish to know the probability distributions that describe the passage of high energy particles through the shock from either side as function of direction (θ_1' and θ_2''). First, the fraction of particles in the ensemble that crosses through the shock-front per unit time from a certain direction $d(\cos\theta)$ scales with the projection of their momenta onto the direction of the crossing (parallel to the shock normal)[5]. $dN_{cross}/dt \propto \hat{\boldsymbol{p}} \cdot (\pm\hat{\boldsymbol{x}}) = \pm\cos\theta$ (down-to-up stream crossing: $+\hat{\boldsymbol{x}}$; up-to-down stream crossing: $-\hat{\boldsymbol{x}}$). Second, on either side of the shock, the presence of turbulent magnetic fields completely randomizes the particle directions and thus the distribution of particle momenta within the ensemble is isotropic: $dN/d\cos\theta \propto 1$. Putting these two considerations together yields (after proper normalization to unity)

$$f\left(\cos\theta_2''\right) = 2 \cdot \cos\theta_2'' \ , \quad 0 \leq \cos\theta_2'' \leq 1 \tag{1.12}$$

$$f\left(\cos\theta_1'\right) = -2 \cdot \cos\theta_1' \ , \quad -1 \leq \cos\theta_1' \leq 0 \tag{1.13}$$

where the support is limited geometrically because only particles with trajectories corresponding to $0° \leq \theta_2'' \leq 90°$ ($90° \leq \theta_1' \leq 180°$) can cross the shock from down (up) stream to up (down) stream, c.f. Fig. 1.3. Thus

$$< \cos\theta_2'' >_{\theta_2''} = \int_0^1 f\left(\cos\theta_2''\right) \times \cos\theta_2'' \ d\left(\cos\theta_2''\right) = \frac{2}{3} \tag{1.14}$$

$$< \cos\theta_1' >_{\theta_1'} = \int_{-1}^0 f\left(\cos\theta_1'\right) \times \left(-\cos\theta_1'\right) \ d\left(\cos\theta_1'\right) = -\frac{2}{3} \tag{1.15}$$

and inserting into eq. (1.11) yields

$$\xi = \frac{1 + \frac{4}{3}\beta + \frac{4}{9}\beta^2}{1-\beta^2} \approx \frac{4}{3}\beta + O\left(\beta^2\right) \tag{1.16}$$

$$\xi \approx \frac{4\left(u_1 - u_2\right)}{3c} + O\left(\left(\frac{u_1 - u_2}{c}\right)^2\right) \tag{1.17}$$

[5] $dN_{cross}/dt \propto \int_A \mathbf{dA} \cdot \mathbf{p}$

where the last approximation follows from Taylor expanding to first order in $\beta = \frac{v}{c}$, justified by the non-relativistic nature of the shock ($\beta \ll 1$). Thus the average fractional gain in energy for one cycle is related to the discontinuous jump of the plasma velocities across the shock. Because this process has a first-order dependence on β, it is sometimes called *first-order Fermi acceleration*[6].

We can predict the energy spectrum of the population of high energetic particles that undergo this type of acceleration process. A particle starting with energy E_0 will, after n cycles, achieve an energy of

$$E_n = E_0 \left(1 + \xi\right)^n \tag{1.18}$$

where the number of cycles necessary to achieve energy E is

$$n = \frac{\log\left(E/E_0\right)}{\log\left(1+\xi\right)} \tag{1.19}$$

However there is always a finite probability $0 < P_{esc} < 1$ that the particle leaves the shock front and thus the acceleration region, hence aborting the sequence. The probability to remain after n encounters is $q \equiv (1 - P_{esc})^n < 1$. Assuming after some (large) time a large number of particles N_p, all injected with energy E_0, remains in the acceleration region, we can find the number of particles $N\left(E_p \geq E\right)$ which experienced at least n cycles from[7]

$$N\left(E_p \geq E\right) = N_p \frac{\sum_{k=n}^{\infty} q^k}{\sum_{k=0}^{\infty} q^k} = N_p q^n \tag{1.20}$$

Inserting eq. (1.19) we find a power-law (integral) spectrum[8]

$$N\left(E_p \geq E\right) = N_p \left(\frac{E_p}{E_0}\right)^{1-\gamma} \tag{1.21}$$

$$\gamma = 1 - \log\left(1 - P_{esc}\right) / \log\left(1+\xi\right) \approx 1 + \frac{P_{esc}}{\xi} \tag{1.22}$$

which yields a power-law differential spectrum[9] with normalization N_p and *spectral index* γ.

$$\frac{dN}{dE} = -\frac{d}{dE} N\left(E_p \geq E\right) = N_p \frac{\gamma - 1}{E_0} \left(\frac{E}{E_0}\right)^{-\gamma} \tag{1.23}$$

To determine the spectral index we need to calculate the escape probability P_{esc}. The flux of relativistic particles moving across the shock from the up-stream side and starting

[6] The fractional energy gain for Fermi's original proposal can be shown to only have $O\left(\beta^2\right)$-dependence, and thus is significantly less efficient [19].

[7] using (partial) geometric series

[8] using $a^{\log b} = b^{\log a}$

[9] Normalize to unity by dividing out N_p. Then this is a *pareto probability distribution* defined as $f\left(x;\alpha\right) = \frac{\alpha x_m^{\alpha}}{x^{\alpha+1}}$, $x \geq x_m$ with $\alpha = \gamma - 1$

a cycle is

$$-\int_{-1}^{0} d\cos\theta_1' \int_{0}^{2\pi} \frac{c\rho_{CR}}{4\pi} \cos\theta_1' = \frac{c\rho_{CR}}{4} \tag{1.24}$$

where we define ρ_{CR} as the number density of the high energy particles and used that the flux is isotropic in the rest frame of the plasma (upstream). The flux of particles convecting away from the shock is given by the down-stream plasma velocity: $\rho_{CR}u_2$. The escape probability thus is the ratio between the two:

$$P_{esc} = \frac{\rho_{CR}u_2}{\frac{1}{4}c\rho_{CR}} = 4\frac{u_2}{c} \tag{1.25}$$

Inserting eqs. (1.25) and (1.11) into eq. (1.22) yields a prediction for the differential spectral index from diffuse shock acceleration of

$$\gamma = 1 + \frac{3}{u_1/u_2 - 1} \tag{1.26}$$

The compression ratio of the shock u_1/u_2 is given by eq. (1.6) and thus one finds for a monatomic gas ($r_c = 5/3$) undergoing a strong shock $M \gg 1$

$$\gamma \approx 2 + \frac{4}{M^2} + O(\frac{1}{M^3}) \longrightarrow \gamma \approx 2 + O(\frac{1}{M^2}) \tag{1.27}$$

The universal prediction from first order Fermi acceleration are high energy particle spectra that behave as $dN/dE \propto E^{-2}$. Insufficiencies in the acceleration process, for example from back-reactions of the accelerated particles on the shock itself, can lead to spectra that are softer $\gamma > 2$ [40]. In particular, γ-ray observations of bright super nova remnants suggest softer high energy spectra with typical spectral indices in the range $\sim [2.2 - 2.4]$ (if interpreted as hadronic in origin) [19][41][42]. Departures in the opposite direction are also possible: amplification of magnetic fields, especially in the early phase of the expanding shock, can produce spectra with $\gamma < 2.0$ [16][43]. Other authors find such scenarios to still permit spectral indices as soft as $\gamma \sim 2.3$ [41].

The observed (galactic) cosmic-ray spectrum at Earth ($\gamma \approx 2.7$) is significantly steeper than the cosmic-ray injection spectrum at the source predicted by simple DSA ($\gamma \approx 2.0$). This can be understood as a propagation effect. Cosmic-ray propagation models predict a rigidity-dependent escape length $\lambda_{esc} \sim R^{-\delta}$ and thus higher-energy cosmic rays are expected to be more efficiently removed from the galaxy. For protons one can show that the observed spectrum at Earth relates to the injection spectrum via $\Phi_{earth} \sim \Phi_s \times E^{-\delta}$. Early measurements of the Boron to Carbon ratio in the cosmic-ray flux at earth indicate $\delta \sim 0.6$. Thus one might infer a (galactic) proton injection spectrum with $\gamma \sim 2.1\,(= 2.7 - 0.6)$, close to the DSA prediction. More recent measurements obtained by PAMELA [44] and AMS-02 [45], however prefer smaller values. The latter measured $\delta = 0.3(3)$. Theoretical predictions based on models of turbulence in magnetized plasmas range from $\delta = 1/2$ [46] to $\delta = 1/3$ [47]. Thus in these scenarios softer proton injection spectra $\gamma \sim [2.2 - 2.4]$ would be deduced.

Similar to their galactic counterparts, the observed spectrum at Earth of UHECRs differs from the injection spectrum at their sources. Interactions between protons and photons from the CMB (cosmic microwave background) as well as infrared photons from the EBL (extra-galactic background light) lead to rapid energy loss over the large distances involved and are responsible for the *GZK cutoff* (c.f. Sec. 1.5). Photo-disintegration of UHECR nuclei increases the energy loss rate for heavier elements [48]. Since the energy loss rates differ between protons and heavier elements, any inference based on the observed cosmic-ray spectrum about the source injection spectrum depends on assumptions about the composition of the injected UHECRs. Pure proton injection would yield soft injection spectra $\gamma > 2.5$ above 10^{18} eV [49] if the maximum energy of the acceleration process reaches $\sim 10^{21}$ eV. If the maximum energy is lower, the injected spectral index could be harder $\gamma \sim 2$ [49]. Recent analysis of data obtained by AUGER [9][50] points instead to a rather heavy composition and is found to require much harder injection spectra: $\gamma \sim 1.0$ [51]. Whether or not the composition at highest energies is dominated by light or heavy nuclei is still debated [52][53].

1.3 Neutrino Production Mechanisms

The neutrinos that we will study in this work stem predominantly from the decay of charged mesons, especially pions.

$$\begin{aligned} \pi^+ &\longrightarrow \nu_\mu + \mu^+ \longrightarrow \nu_\mu + e^+ + \nu_e + \bar{\nu}_\mu \\ \pi^- &\longrightarrow \bar{\nu}_\mu + \mu^- \longrightarrow \bar{\nu}_\mu + e^- + \bar{\nu}_e + \nu_\mu \end{aligned} \tag{1.28}$$

These pions are created when high-energy protons interact and thus neutrino production can take place wherever such interactions take place. For this reason every astrophysical, hadronic particle accelerator is expected to be a source of neutrinos. While protons are confined by magnetic fields during acceleration, the neutrinos are not and thus can freely escape the source once produced. However the production is not limited to the accelerator itself. The high-energy protons that do escape from the source can interact with gas (e.g. molecular clouds) on their path and produce neutrinos. We refer to neutrinos that are not produced at Earth as *astrophysical neutrinos*. Finally, once at Earth, interactions inside Earth's atmosphere produce *atmospheric neutrinos* via the same mechanism (c.f. Sec. 1.7).

Pion production giving rise to a flux of astrophysical neutrinos can broadly be classified into two mechanisms: purely hadronic scenarios dominated by proton-proton (pp) collisions and photo-hadronic ($p\gamma$) pion production due to protons interacting with ambient photons (photons created by the source itself, but also CMB photons, EBL photons). These two scenarios can be seen as limiting simplifications of more complex source models [54]. For pp-dominated sources one expects

$$p + p \longrightarrow N\left[\pi^0 + \pi^+ + \pi^-\right] + X \tag{1.29}$$

while for $p\gamma$-dominated sources one finds

$$p + \gamma \longrightarrow \Delta^+ \longrightarrow \begin{cases} \pi^+ + n & \frac{1}{3} \text{ of all cases} \\ \pi^0 + p & \frac{2}{3} \text{ of all cases} \end{cases} \tag{1.30}$$

In both scenarios the decays of the neutral prions $\pi^0 \to \gamma + \gamma$ give rise to a flux of high-energy γ-rays. The implications will be discussed in Sec. 1.6. Neutrinos produced via the pion decay chain carry on average a fraction η of the kinetic energy of the primary proton

$$E_\nu \sim \eta \cdot E_p \quad \text{with} \quad \eta \sim 0.05 \tag{1.31}$$

Eq. (1.30) permits the production of additional (low energy [19]) electron anti-neutrinos from the decay of the neutron. The injected neutrino flavor ratios $(\nu_e : \nu_\mu : \nu_\tau)$ and $(\bar{\nu}_e : \bar{\nu}_\mu : \bar{\nu}_\tau)$ depend on the details of energy-loss mechanisms at the source. The baseline scenario uses eq. (1.29) and predicts $(\nu_e : \nu_\mu : \nu_\tau) = (\bar{\nu}_e : \bar{\nu}_\mu : \bar{\nu}_\tau) = (1 : 2 : 0)$ for pp-dominated sources. Neutrino oscillations over astrophysical distances will alter the injected flavor ratio and one expects to observe a (nearly) equal admixture of neutrino flavors in the measured flux $(\nu_e : \nu_\mu : \nu_\tau) \approx (1 : 1 : 1)$ at Earth[10].

More exotic neutrino production mechanisms include top down scenarios in which heavy dark matter particles decay into high-energy neutrinos [56][57][58].

1.4 Astrophysical Sources of High Energy Neutrinos

Obvious sources of astrophysical neutrinos are the sun (solar neutrinos) and core-collapse supernova explosions (SN neutrinos). Both produce low energy neutrinos[11] as shown in Fig. 1.4. Their energies are below the particle detection threshold of IceCube ($\sim 10\,\text{GeV}$) and are thus of no further interest for this work.

Galactic Supernova Remnants

From the discussion above it should be clear that galactic SNRs are expected to produce a flux of astrophysical neutrinos at energies much larger ($\sim$ TeV) than the ones produced in the actual supernova explosion ($\sim$ MeV). From the perspective of a neutrino detector, these sources can be considered as point-like. SNRs that appear bright in γ-rays with a hard spectrum $\gamma \sim 2$ are especially promising candidate neutrino sources, if their γ-ray emission is rooted in hadronic interactions. One example is SNR RX J1713.7-3946, which has been observed by the HESS γ-ray telescope at energies $E_\gamma > 10\,\text{TeV}$. The observed flux of γ-rays allows to derive an order-of-magnitude estimate of its neutrino emission of $E^2 dN/dE \sim 10^{-11}\,\text{TeV}\,\text{s}^{-1}\,\text{cm}^{-2}$ [16], corresponding to an integral flux of $\sim 2\,\text{km}^{-2}\text{yr}^{-1}$ above 1 TeV [61]. IceCube has searched for neutrino emission from point sources in general including RX J1713.7-3946 (and other galactic SNRs for that matter)

[10] This expectation can be turned around and has been proposed to search for non-standard neutrino oscilllations [55].

[11] An ensemble of SN neutrinos can still be detected by IceCube as a correlated enhancement in the noise-rate of all DOMs across the detector [60].

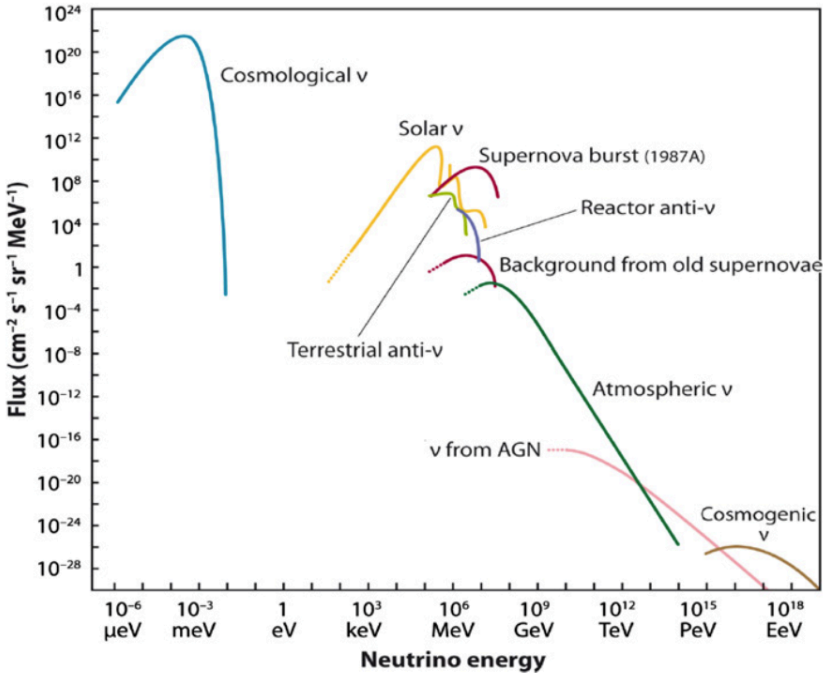


FIGURE 1.4: Comparison of Neutrino Fluxes expected from different natural and man-made neutrino sources. (Figure from [59])

in the muon neutrino channel. Thus far, no neutrino point source has been identified. For RX J1713.7-3946 the current flux upper-limit[12] is similar to this rough estimate [62]. The physics of galactic neutrino emission is reviewed in [63].

Gamma Ray Bursts

Gamma Ray Bursts (GRBs) are powerful flashes of high energy γ-rays, first observed in the late 1960s. They can release a total energy of up to 10^{53} erg within a few seconds [19]. Since the distribution of GRBs across the sky is isotropic, they are extra-galactic in origin. Long GRBs ($t > 2\,\mathrm{s}$) are believed to be due to black hole formation of massive stars releasing the energy in powerful relativistic jets. Such jets could power particle acceleration [19]. Short GRBs ($t < 2\,\mathrm{s}$) are instead believed to be caused by merging neutron star pairs (or neutron star - black hole pairs). The recently observed gravitational wave event GW170817 [6] in spatial and temporal coincidence with a burst of γ-rays detected by the Fermi satellite are consistent with the merger of two neutron stars [64]. While the physics of the GRB emission is not fully understood, it has been suggested that $p\gamma$ interactions between shock-accelerated protons $\sim 10^{15}\,\mathrm{eV}$ with $\sim$ MeV γ-rays could produce a flux of $\sim 10^{14}\,\mathrm{eV}$ neutrinos [19][65]. Therefore, GRBs are also promising candidate sources for UHECRs (with some caveats [66]). GRBs spectra are found to be described by a doubly-broken power law [19]

$$\frac{dN}{dE} \propto \begin{cases} E^{-\alpha}\exp\left(-E/E_0\right), & E < E_0 \\ E^{-\beta}, & E \geq E_0 \end{cases} \tag{1.32}$$

[12] IceCube obtains significantly stronger limits for sources that are located in the Northern Sky.

with $\beta > \alpha$. While all parameters are variable between GRBs, typical values for the spectral indices can be defined ($\alpha \approx 1$, $\beta \approx 2$).
IceCube has searched for neutrino emission from known GRBs but none was found, thus challenging model assumptions [67]. However more recent estimates lowered the neutrino flux expectations [68][69]. We refer to [70] for a recent review of the topic.

Active Galactic Nuclei
Active Galactic Nuclei (AGNs) refer to galaxies with an exceptionally bright central region and make up about $O(1)\%$ of all galaxies [19]. Thus, by definition, AGNs are extra-galactic sources. The emission is powered by a supermassive central blackhole. Gravitational energy is released in the accretion process of material onto the black hole. AGNs have been suggested as possible sources of UHECRs [71]. Historically, AGNs are classified by their observational phenomenology. From the perspective of neutrino emission radio-loud AGNs with relativistic, jetted outflows are of particular interest. If the jet of a radio-loud AGN points towards Earth, the AGN is classified as *blazar*. Blazar jets are good candidates for hadronic particle acceleration via DSA with neutrinos produced in $p\gamma$ interactions. However the observed γ-ray spectra could also be explained with purely leptonic processes [16], in which case no neutrinos would be produced. Despite in jet, neutrinos (and high energy cosmic rays) could also be produced closer to the central engine (black hole) [72][73]. In this case the neutrino emission would not relate to any observable γ-rays, because photons produced near the blackhole, where densities are large, in contrast to the neutrinos would simply get absorbed. Hence neutrinos would be the only proof of hadronic interactions near the black hole [48]. Depending on the distance to the central black-hole AGN jets have the potential to generate neutrinos with energies beyond $\sim 10^{18}\,\mathrm{eV}$ [48]. Early indications of correlations between the arrival directions of UHECRs with known AGNs have since disappeared [9].
An interesting connection between γ-rays, UHECRs and high energy neutrino production has been pointed out in [74]. In this model high energy protons are accelerated in black-hole jets that are embedded in clusters of galaxies. Subsequent interactions with the cluster medium can efficiently produce secondary neutrinos and γ-rays, broadly consistent with observations in all three channels [74].

Starburst Galaxies
Starburst galaxies (SBGs) are yet another potential extra-galactic source of high energy neutrinos. These are galaxies with large gas densities that allow for exceptionally high specific star formation rates (sSFR), star formation rate per galaxy stellar mass, larger than the average sSFR of galaxies at a given redshift z [75]. Cosmic rays produced in such galaxies (e.g. in SNRs) would lose all their energy in interactions with the ambient gas. The resulting charged and neutral pions from e.g. pp-collisions produce γ-ray and neutrino spectra that follow the cosmic-ray injection spectra more closely than in our galaxy, where leakage of high energy cosmic rays from the galaxy has to be considered (c.f. Sec. 1.2). The cumulative neutrino flux from SBGs has been estimated in [76]:

$$E^2 \frac{dN}{dE} \approx 10^{-7} \left(\frac{E}{1\,\mathrm{GeV}} \right)^{-0.15 \pm 0.1} \mathrm{GeV\,cm^{-2}\,s^{-1}\,sr^{-1}} \tag{1.33}$$

thus the spectral index ranges from $2.0 \lesssim \gamma \lesssim 2.3$. At high energies $E > 100\,\mathrm{TeV}$, the morphology of the flux might be more complicated. As in our galaxy, a possible knee

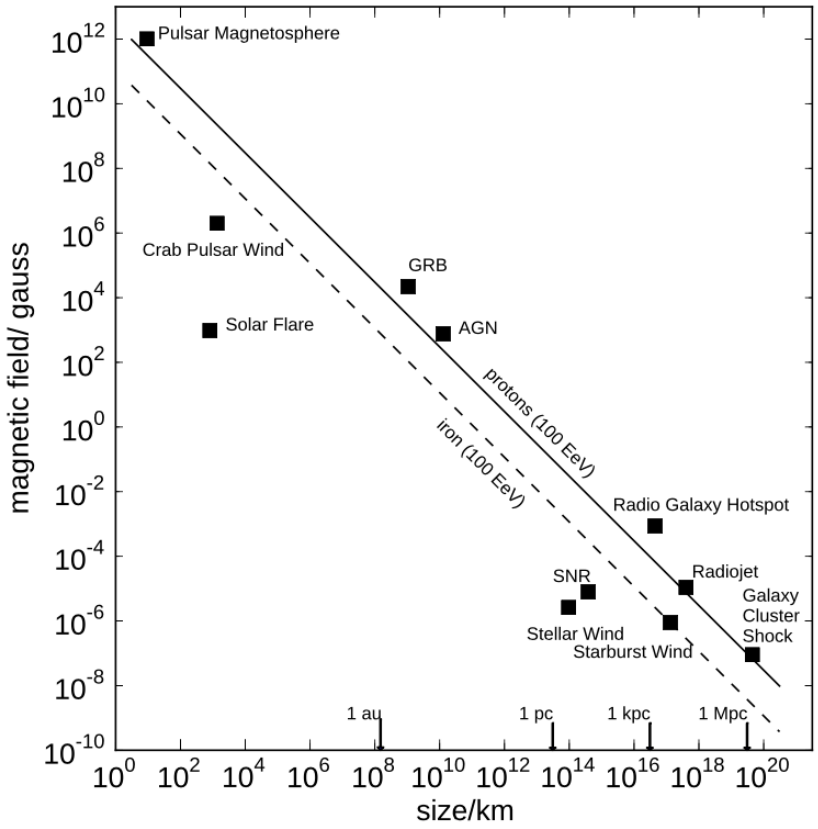


FIGURE 1.5: The Hillas criterion eq. (1.4) evaluated for various high energy cosmic ray candidate sources. (Figure from [80])

in the spectrum of primary cosmic rays would introduce a break into the spectrum of SBG neutrinos. The onset of the knee in SBGs, however, might be shifted towards higher energies, given that SBGs are believed to have higher magnetic fields than out own galaxy. Other authors have pointed out that $\gamma < 2.2$ is imposed by measurements of the diffuse flux of γ-rays by Fermi [77], that would otherwise be over-predicted.
Most recently it has been pointed out [78][79] that fluxes as high as predicted by eq. (1.33) are strongly challenged by recent Fermi measurements, that attribute a large fraction of the observed diffuse γ-ray flux to non-identified blazars, i.e. blazars too faint to be identified individually by Fermi (c.f. Sec. 1.6).

The maximum neutrino energies reachable by different sources are related to the limiting proton energy via eq. (1.31), which can be estimated from the Hillas criterion eq. (1.4). In this simple picture, the maximum proton energy depends on the magnetic field and the size of the source. These are compared in Fig. 1.5 ("Hillas plot") for a variety of sources, including the ones discussed above.

An Upper Bound on the Extra Galactic Neutrino Flux
Assuming that extra galactic neutrinos are produced in sources that are optically thin to cosmic rays (i.e. where cosmic rays can escape from the source alongside the neutrinos) it is possible to derive constraints on the expected neutrino flux from such sources using

UHECR measurements [81][82][83]. Here we follow the argument presented first in [81], known as the *Waxman-Bahcall bound*, which assumes a CR injection spectrum with spectral index $\gamma = 2$ (from DSA, c.f. Sec. 1.2) and with proton-dominated composition. The observed flux of UHECR cosmic rays ($E > 10^{19.2}\,\mathrm{GeV}$) allows for an estimation of the local ($z = 0$) generation rate of cosmic-ray energy (per logarithmic energy interval) of [75]

$$dQ/d\ln E \equiv \left(E_p^2 \frac{d\dot{n}_p}{dE_p}\right)_{z=0} = (0.5 \pm 0.15) \times 10^{44}\,\mathrm{erg\,Mpc^{-3}\,y^{-1}} \tag{1.34}$$

where z is the redshift parameter. In the vicinity of the source, protons have some probability to undergo photo-meson production interactions, eq. 1.30. Thus a fraction ϵ of these protons will convert their energy into the production of secondary particles. The resulting neutrinos (from pion decay, eq. (1.28)) inherit the injected proton spectrum, since in each interaction a part $\eta \sim 0.05$ of the parent proton energy is transferred to the neutrino (see eq. (1.31)). These considerations make it possible to relate the expected neutrino flux to the generated cosmic-ray energy $dQ/d\ln E$

$$E_\nu^2 \frac{dN_\nu}{dE_\nu} \approx \frac{1}{4}\,\xi_z t_H \frac{c}{4\pi} \times \epsilon \frac{dQ_{CR}}{d\ln E_{CR}} \tag{1.35}$$

where $t_H \sim 10^{10}\,\mathrm{y}$ is the Hubble time and ξ_z a redshift correction. The factor of $1/4$ accounts for the fact that on average per photo-meson interaction, eq. (1.30), mostly γ-rays are produced and thus only a fraction of the non-hadronic secondary particles are neutrinos. The upper limit on the neutrino flux is then given by $\epsilon \to 1$, when all protons undergo one photo-meson production interaction. In this limit cosmic-rays will still be observable, because neutrons are produced (e.g. eq. (1.30)) that can escape the source and decay later into protons [81].
Numerically, the limiting all-flavor neutrino flux is given as [75]

$$E_\nu^2 \frac{dN}{dE_\nu}_{WB} = 3.4 \cdot 10^{-8} \times \frac{\xi_z}{3} \left[\frac{dQ_{CR}/d\ln E_{CR}}{0.5 \cdot 10^{44}\,\mathrm{erg\,Mpc^{-3}\,y^{-1}}}\right] \mathrm{GeV\,cm^{-2}s^{-1}sr^{-1}} \tag{1.36}$$

The redshift correction ξ_z accounts for the possibility that the cosmic-ray generation rate could have been different (higher) in the cosmological past. Such protons can not reach Earth because of the GZK effect (c.f. 1.5) but the neutrinos would, albeit with reduced energy (red-shifted). Values for ξ_z range from $\xi_z = 0.6$ (no evolution) to $\xi_z = 3$ (evolution similar to star formation rate). Thus, the conservative bound becomes

$$E_\nu^2 \frac{dN}{dE_\nu}_{WB} \lesssim 3 \cdot 10^{-8} \times \left[\frac{dQ_{CR}/d\ln E_{CR}}{0.5 \cdot 10^{44}\,\mathrm{erg\,Mpc^{-3}\,y^{-1}}}\right] \mathrm{GeV\,cm^{-2}s^{-1}sr^{-1}} \tag{1.37}$$

The bound does not apply to sources where no cosmic-rays can escape, for example due to large optical thickness of the source to proton-nucleon interactions. Possible models are given in [72][73]. Finally higher neutrino fluxes are allowed if the proton injection spectrum is softer than $\gamma = 2$.

1.5 Neutrinos from Cosmic Ray Propagation

Galactic Propagation
During propagation from their sources, cosmic rays can produce neutrinos in pp-collisions with interstellar gas [39].

$$\frac{dN_\nu}{dE_\nu} \sim \tau_\nu\left(E_\nu\right) \frac{dN_p}{dE_p}\Big|_{E_p=E_\nu} \tag{1.38}$$

where $\tau_\nu\left(E_\nu\right)$ is an effective optical depth for neutrino production and depends on the gas and cosmic ray densities within our galaxy as well as the proton interaction cross-section. Here one has to assume that the cosmic-ray flux observed at Earth is representative of the flux throughout the galaxy, which is subject to debate[84][85]. Using eq. (1.38) the neutrino flux from galactic cosmic ray propagation has been approximated in [39]

$$\frac{dN_\nu}{dE_\nu} \lesssim 4 \times 10^{-22} \left(\frac{E^2 dN_p/dE_p\Big|_{E=1\,\mathrm{PeV}}}{3 \cdot 10^{-4}\,\mathrm{GeV\,cm^{-2}\,s^{-1}\,sr^{-1}}}\right) \times \left(\frac{E_\nu}{1\,\mathrm{PeV}}\right)^{-2.7} \tag{1.39}$$
$$\mathrm{GeV^{-1}cm^{-2}s^{-1}sr^{-1}}$$

which at $E_\nu = 1\,\mathrm{PeV}$ would be $O(1)\%$ of the Waxman-Bahcall bound, eq. (1.36). For comparison: IceCube observes a flux of astrophysical neutrinos at $\sim 1\,\mathrm{PeV}$ at the level of the WB-bound (eq. (1.37)). More refined models for diffuse galactic neutrino production based on γ-ray observations and simulations of galactic cosmic ray propagation suggest higher neutrino fluxes [86]. Using these models, IceCube has searched for such a neutrino flux above $E_\nu = 1\,\mathrm{TeV}$. The non-observations thereof allows to constrain the galactic contribution to the diffuse astrophysical neutrino flux, measured by IceCube, to $< 14\%$ [87].

Extra Galactic Propagation
Cosmic ray protons (and nuclei) in the UHECR energy regime rapidly lose energy in interactions with photons (CMB and EBL). The interactions include [88] the production of electron-positron pairs ($p + \gamma \rightarrow p + e^+e^-$), photo-pion production eq. (1.30) and photo-disintegration of nuclei (e.g. $Z^A + \gamma \rightarrow Z^{A-1} + n$). In all interactions energy is transferred from the proton (nuclei) to the secondary particles. For pair-production the proton energy threshold is [39]

$$E_p \sim 4.8 \times 10^{17} \left(\frac{\epsilon}{10^{-3}\,\mathrm{eV}}\right)^{-1} \mathrm{eV} \tag{1.40}$$

where ϵ_γ is the energy of the target photon. The photo-pion production threshold is higher since the (rest) mass of the pion is higher than that of the electron-positron pair:

$$E_p \sim 3.4 \times 10^{19} \left(\frac{\epsilon_\gamma}{10^{-3}\,\mathrm{eV}}\right)^{-1} \mathrm{eV} \tag{1.41}$$

The corresponding energy loss lengths $dE/ds = -E/L_{loss}$ for high energy protons (nuclei) are shown in Fig. 1.16, left (right). The energy loss length of protons (nuclei) at highest energies ($\sim 10^{20}\,\mathrm{eV}$) is small compared to the observable universe. Taking into account

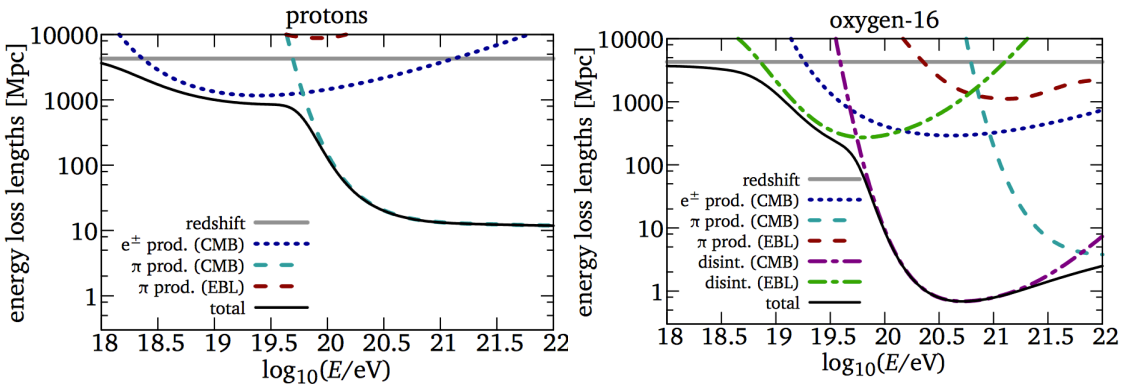


FIGURE 1.6: Energy loss length of UHECRs as function of primary energy for protons (left) and oxygen (right) for different energy loss mechanisms (Figure from [88]).

cosmological effects, one can show that the cosmic rays observed at Earth, if interpreted as protons, must have been produced within the *GZK sphere* [89][90] of radius $75-150\,\mathrm{Mpc}$ [19]. Similar arguments apply to heavy primaries [91]. Whether or not the observed suppression of the flux of UHECR cosmic rays above $\sim 40\,\mathrm{EeV}$ is due to this GZK cutoff or simply related to the maximum energy, that the UHECR sources can accelerate particles to, remains to be seen [92].

If protons dominate the cosmic-ray spectrum at highest energies and their sources are powerful enough to accelerate them to energies beyond the GZK cutoff, then neutrinos are generated efficiently through the pion decay chain eq. (1.28). The corresponding flux of *GZK neutrinos* could be observed at Earth. Flux estimates are strongly model dependent [93]. IceCube has searched for this neutrino flux but so far no neutrinos consistent with a possible GZK origin have been observed[13] [94].

1.6 Relevance of γ-ray Observations

Neutrino production at extra-galactic sources (pp or $p\gamma$) is associated with the production of γ-rays from the decay of neutral pions, see eq. (1.29) and eq. (1.30). Assuming that these sources are optically thin to photons, such that the γ-rays can escape instead of being absorbed, it is possible to constrain the diffuse astrophysical neutrino flux and possible sources thereof. During propagation, these high energy γ-rays are subject to interactions, similar to the case of high energy protons. Of particular relevance is pair-production on background photons from the CMB: $\gamma + \gamma_{CMB} \longrightarrow e^+e^-$ (but also other cosmic photon radiation, e .g. EBL). The energy threshold for high-energy photons (E_γ) to produce an electron-positron pair with a background photon of energy ϵ_γ is [39]

$$E_\gamma \sim 2.6 \times 10^{14} \left(\frac{\epsilon_\gamma}{10^{-3}\,\mathrm{eV}}\right)\,\mathrm{eV} \tag{1.42}$$

[13] From $E_\nu \sim 0.05 E_p$ one expects such neutrinos to have $O(EeV)$ energies

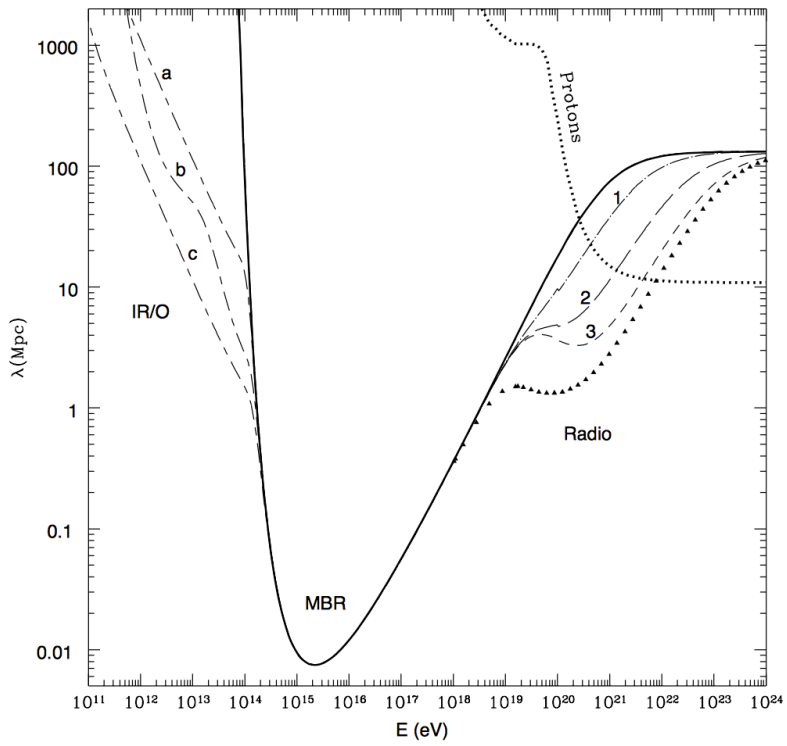


FIGURE 1.7: Mean free path of very high energy photons due to pair production interactions with low energy photons from CMB (solid line) and other cosmic photon radiation (dashed, dash-dotted). (Figure from [95])

and thus photons of $O(300)$ TeV are efficiently absorbed. However through inverse compton scattering ($e^{\pm} + \gamma_{CMB} \longrightarrow \gamma' + e^{\pm}$) it is possible for one of the high energy leptons to create another high energy photon ($E'_{\gamma} < E_{\gamma}$). The produce repeats until the leading photon energy falls below the pair-production threshold and the cascade stops. The low-energy photons at the end of the cascade can be measured by γ-ray satellites, like Fermi [32]. The constraint on the diffuse astrophysical neutrino flux is derived by requiring that the expected diffuse γ-ray flux from this cascading process not exceed the measured flux of γ-rays. Such a calculation has been performed to formulate a *cascade bound* on the expected flux of GZK-neutrinos (see above) [96]. It is similar to the Waxman-Bahcall bound and is given in for the sum of all flavors [96].

$$E_{\nu}^2 \frac{dN}{dE_{\nu}} \lesssim 4 \times 10^{-8}\,\mathrm{GeV\,cm^{-2}\,s^{-1}\,sr^{-1}} \tag{1.43}$$

Based on similar arguments it has recently been pointed out, that the astrophysical neutrino flux measured by IceCube in the few tens of TeV energy range, appears to favor sources from which γ-rays do not escape [78][79].
The mean free path of high-energy photons as function of their energy (and photon target) is given in Fig. 1.15. The strong high energy absorption of γ-rays is one of the motivations for pointed neutrino astronomy. The universe at highest energies is opaque to photons but not neutrinos.

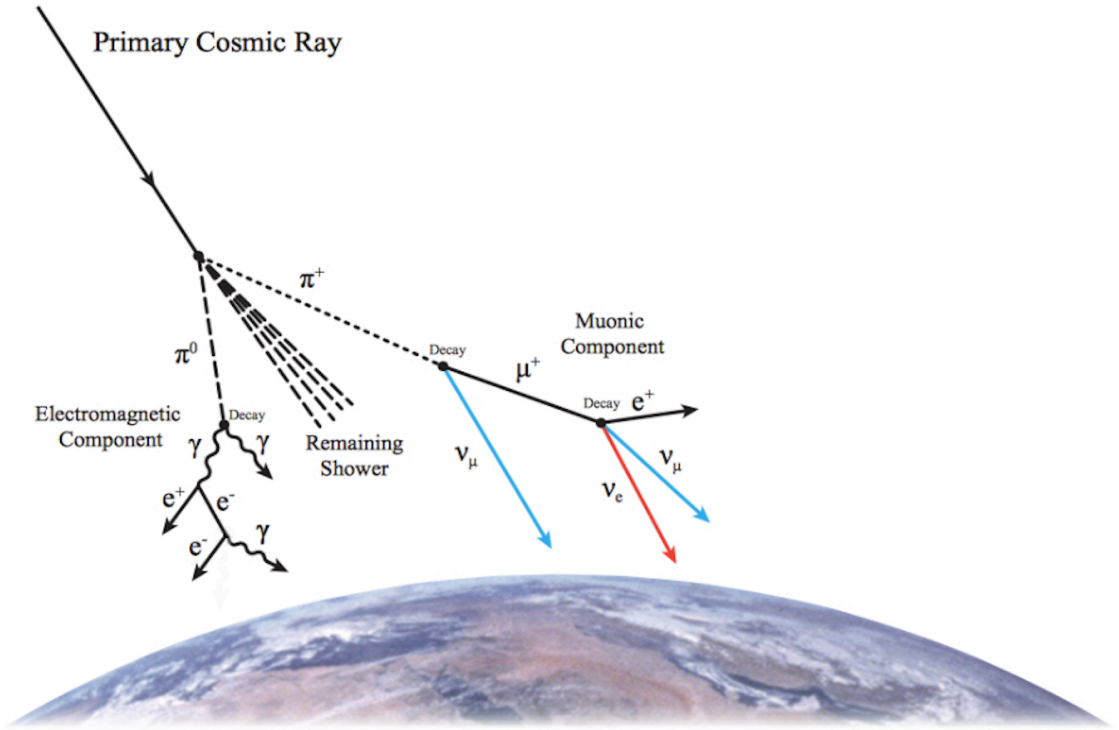


FIGURE 1.8: A schematic diagram of secondary particle production during the development of cosmic-ray induced air-showers. (Figure from [97])

1.7 Atmospheric Neutrinos

Cosmic Rays not only produce neutrinos in astrophysical sources or during propagation but also upon arrival at Earth. Nucleon-nucleon collisions between cosmic ray nuclei and nuclei in Earth's atmosphere initiate cascades of secondary particles, called air showers, that develop towards the surface. Neutrinos are produced in the decays of mesons: pions, kaons and, to a lesser degree, heavier mesons. This is schematically visualized in Fig. 1.8. The fluxes of secondary particles at the surface depend on the competition between interaction and decay of unstable particles in the shower. Qualitatively higher-energy pions and Kaons are more likely to interact before they decay and thus on average suffer from higher energy loss than pions and Kaons of lower energy. Neutrinos that stem from the decay of pions and kaons in the atmosphere are called *conventional atmospheric neutrinos*, while neutrinos produced in the decays of heavier mesons (involving a charm quark) are referred to as *prompt atmospheric neutrinos*. The latter name reflects that the lifetime of heavy mesons is so short that interactions (thus energy loss) are negligible and hence the energy spectrum at Earth follows that of the primary cosmic-rays $\sim E^{-2.7}$. The conventional component instead follows the primary cosmic-ray spectrum only at low energies $\ll 150\,\mathrm{GeV}$. At higher energies $\gg 850\,\mathrm{GeV}$ energy loss steepens the spectrum by about one power compared to the CRs and behaves as $\sim E^{-3.7}$. The meson decay probabilities also influence the declination dependence of the two atmospheric neutrino flux components. Prompt atmospheric neutrinos again inherit their flux from that of the primary cosmic-rays, and thus their flux is isotropic. The conventional atmospheric neutrino flux is largest for horizontal directions, because the effective depth of the atmosphere

is larger, increasing the fraction of mesons that decay rather than interact compared to vertical directions. At energies relevant for this work and above, $E \gtrsim 1\,\mathrm{TeV}$, the conventional atmospheric flux is increasingly dominated by the contribution from the decays of charged Kaons into muon neutrinos. The prompt atmospheric neutrino flux instead consists equally of muon and electron neutrinos, with negligible contribution from tau neutrinos.

Quantitatively the evolution of secondary particle fluxes during air-shower development are described by a set of coupled differential equations, called *cascade equations* [98]. The flux Φ_h of secondary particles of type h changes within a differential slant depth dX as

$$\begin{aligned}\frac{d}{dX}\Phi_h\left(E\right) = &-\left\{\frac{\Phi_h\left(E\right)}{\lambda^h_{int}\left(E\right)} + \frac{\Phi_h\left(E\right)}{\lambda^h_{dec}\left(E\right)}\right\} \\ &+ \sum_l \int_{E'>E} dE' \left\{\frac{1}{\lambda^l_{int}\left(E'\right)}\frac{C_{l(E')\to h(E)}}{E} + \frac{1}{\lambda^l_{dec}\left(E'\right)}\frac{D_{l(E')\to h(E)}}{E}\right\}\Phi_l\left(E'\right)\end{aligned} \tag{1.44}$$

where the slant depth $X\left(h_0\right) = \int_0^{h_0} dl \rho_{air}\left(h_{atm}\left(l\right)\right)$ is defined along the particle trajectory l. The first term in eq. (1.44) (RHS) describes the loss of particles of type h with energy E through interactions and decay, while the second term gives the contribution to the flux of h at energy E from interactions and decay of all other particles in the cascade with energies $E' > E$. A discretized version of eq. (1.44) can be solved numerically using the *Matrix Cascade Equation* (MCEq) package [98].

Alternatively, one can calculate the atmospheric neutrino flux at surface from Monte Carlo simulations of the particle transport during air-shower development. Such simulations have been performed by various groups [99][100].

Approximate analytical solutions are given in [19] for the conventional neutrino flux (π/K-component)

$$\begin{aligned}\frac{dN_\nu}{dE_\nu}\left(e_\nu\right) \approx \frac{dN_{N_{CR}}}{dE_{N_{CR}}}\frac{1}{1-Z_{NN}}\left(E_\nu\right) \times \Bigg\{&\frac{A_{\pi\nu}}{1 + B_{\pi\nu}\cos\theta E_\nu/\epsilon_\pi} \\ &+ 0.635\frac{A_{K\nu}}{1 + B_{K\nu}\cos\theta E_\nu/\epsilon_K}\Bigg\}\end{aligned} \tag{1.45}$$

where $dN_{N_{CR}}/dE_{N_{CR}}$ is the primary cosmic ray nucleon flux. The branching ratio for the leptonic $K^\pm$ decay channel is 0.635 and Z_{NN} is a spectrum weighted moment for nucleon (p or n) production in cosmic-ray collisions with air[14]. Similarly, the coefficients A_{ij} and B_{ij} encode particle physics relevant for pion and kaon production in cosmic ray-air collisions. Lastly, the critical energies $\epsilon_{\pi/K}$ ($\epsilon_\pi = 115\,\mathrm{GeV}$, $\epsilon_{K^\pm} = 850\,\mathrm{GeV}$ [19]) denote the energy above which interactions dominate over meson decay and explain the steepening of the conventional neutrino spectrum compared to the primary cosmic rays for $E_\nu \gg \epsilon_{\pi/K}$. For near horizontal trajectories $\theta \gtrsim 70°$ the zenith angle $\cos\theta$ in eq. (1.45) need to be corrected for the curvature of atmosphere. This correction is given in [103].

[14] This is essentially the average fraction of the interaction energy going into nucleon final states (p or n) [19].

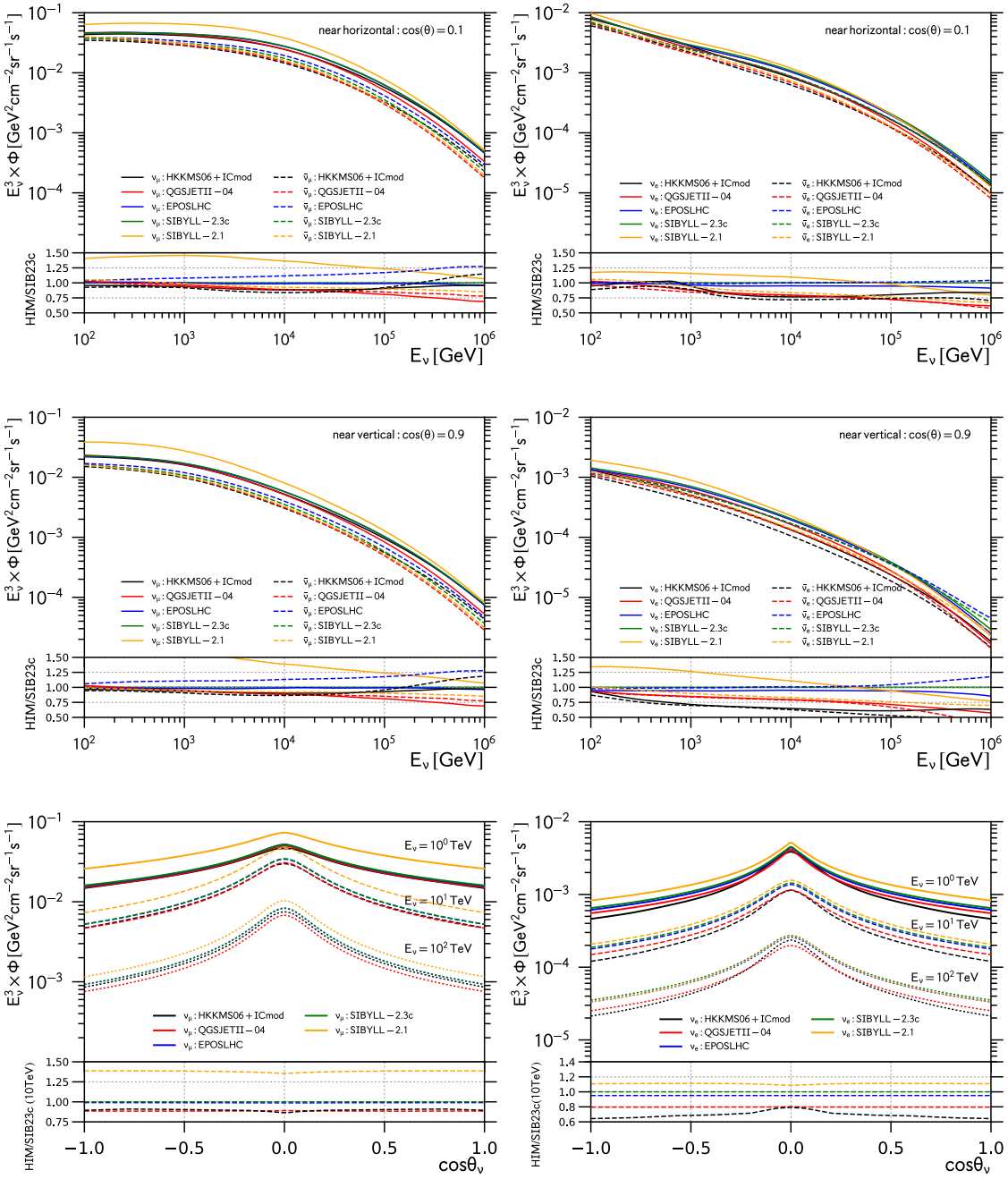


FIGURE 1.9: Conventional atmospheric muon neutrino (left) and electron neutrino (right) flux predictions for various hadronic interaction models (colors). Top (center) row: Energy spectrum for nearly horizontal (vertical) directions. Bottom row: flux as function of zenith angle at different energies. Flux predictions from [99] (black) and [98] (other colors).

The nominal conventional atmospheric flux model, used by IceCube, relies on the Monte Carlo simulation of [99] (HKKMS06), which above $E_\nu = 10\,\text{TeV}$ has been extrapolated to higher energies using eq. (1.45). The primary cosmic ray flux assumes the Gaisser-H3a model, eq. (1.45)[15]. The corresponding conventional atmospheric neutrino flux prediction is shown in Fig. 1.9 (black) for muon neutrinos (left) and electron neutrinos (right). The flux is sensitive to assumptions about hadronic interactions in the atmosphere.

As part of our work, we used MCEq [98] to numerically solve the cascade equations for different interactions models to derive an independent flux prediction. The corresponding solutions are given in Fig. 1.9 (colors). The differences constitute systematic uncertainties and are discussed in Sec. 4.2.1. Since IceCube is not embedded within a strong magnetic field, the detector essentially has no capabilities to distinguish between neutrinos and

[15]Since the original MC simulation performed by HKKMS06 assumed a different (dated) cosmic-ray primary flux, a correction function is applied.

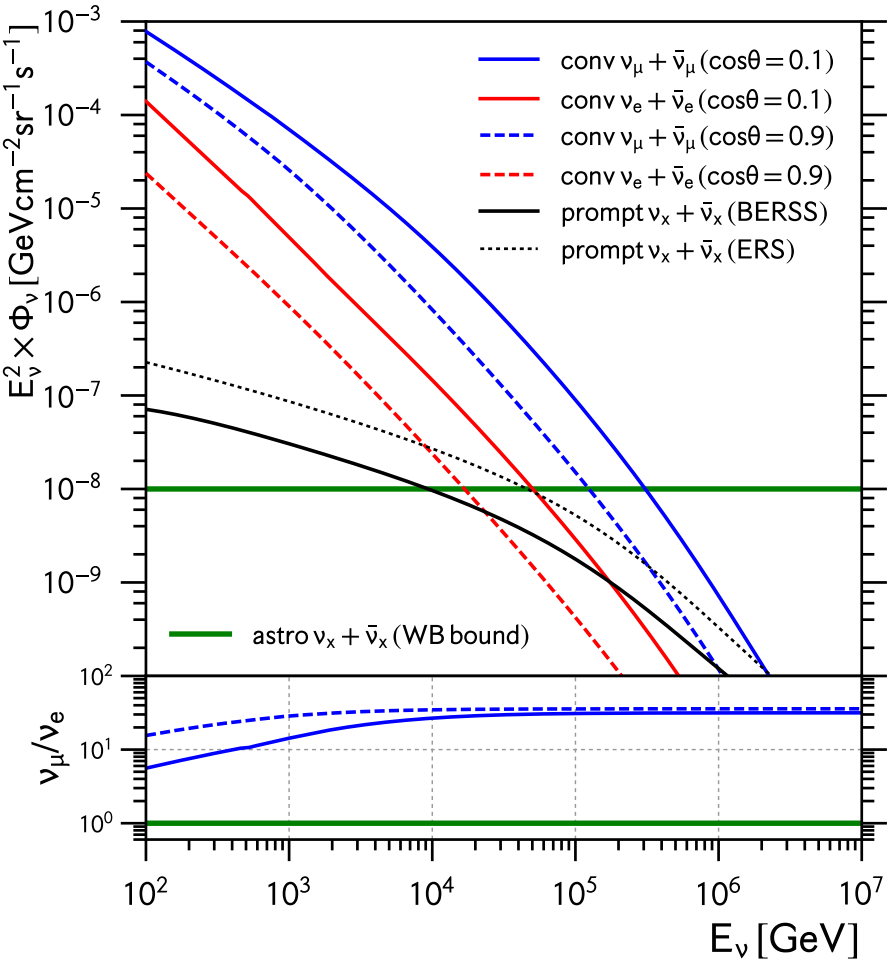


FIGURE 1.10: Neutrino Fluxes relevant for this work: conventional atmospheric muon neutrinos (blue) and electron neutrinos (red) [99]; prompt atmospheric neutrinos (black) [101][102] and Waxman-Bahcall bound [81] for astrophysical neutrinos (green).

anti-neutrinos[16] and thus it is instructive to discuss fluxes per neutrino flavor ($\nu_x + \bar{\nu}_x$, $x \in \{e, \mu, \tau\}$)[17]. Since the tau lepton is heavier than pions and kaons, the conventional neutrino flux does not contain a tau neutrino contribution. Fig. 1.10 compares the total conventional flux of electron neutrinos (red) to that of muon neutrinos (blue). Above $E_\nu \sim 10\,\text{TeV}$ the contribution from muon neutrinos dominates over that from electron neutrinos by a factor of ~ 30, since electron neutrinos are only produced in semi-leptonic Kaon ($K^\pm$, K^0_L)[18] decays. For comparison the Waxman-Bahcall per-flavor upper bound is shown in green.

Prompt Atmospheric Neutrinos

Prompt atmospheric neutrinos are created in the decays of heavy mesons including at least a charm (or heavier) quark that are produced in energetic air-showers. The flux prediction suffers from larger uncertainties than that of conventional atmospheric neutrinos, since high-energy forward production of heavy mesons is experimentally not well constrained. Thus far, the prompt neutrino flux has not been observed. In the past, IceCube analyses used the ERS [101] prompt neutrino flux prediction. The calculation has later been refined by the same authors [102], which we will refer to as BERSS prompt neutrino flux prediction. At relevant energies $1\,\text{TeV} < E_\nu < 1\,\text{PeV}$ the BERSS flux is

[16] up to small differences in inelasticity distributions.

[17] At low energies νN and $\bar{\nu}$N DIS interaction cross-sections are different c.f. Sec, 1.8.

[18] At highest energies $\sim 100\,\text{TeV}$ an extra contribution to the electron neutrino flux from K^0_S decays is expected [42].

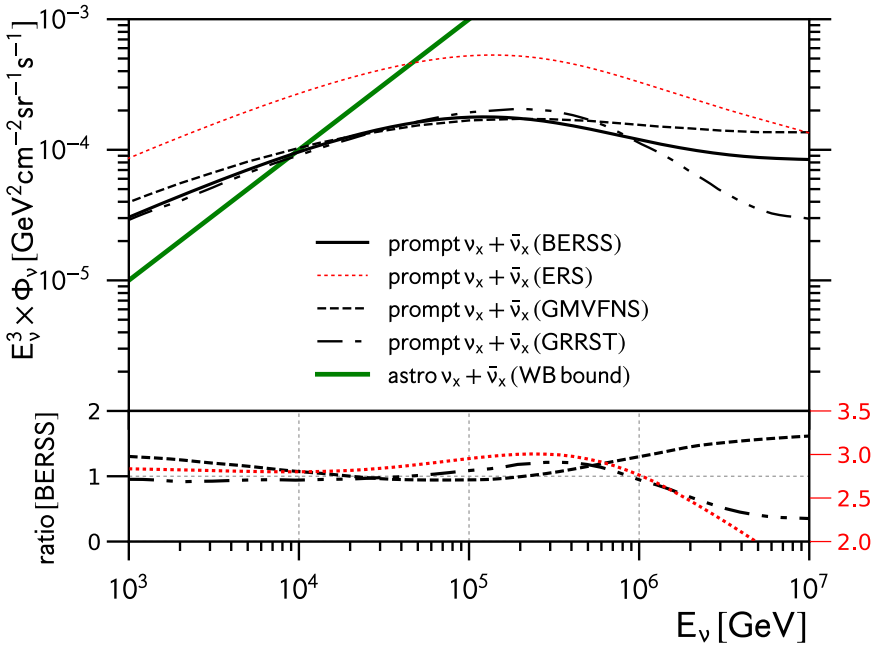


FIGURE 1.11: Various calculations of the atmospheric neutrino flux from the decay of heavy mesons (prompt): ERS [101], BERSS [102], GRRST [104] and GMVFNS [105] in comparison to the WB bound [81] on the flux of astrophysical neutrinos.

lower than the earlier ERS prediction by a factor of $\sim 2.7-3.0$[19]. Both fluxes (per flavor) are compared to the conventional neutrino flux (and the astrophysical WB bound) in Fig. 1.9 (ERS: black dotted, BERSS: black solid), assuming the Gaisser-H3a primary cosmic-ray flux, eq. (1.5). Since the prompt fluxes of electron and muon neutrinos are equal and also independent of declination, only one line is visible. Not shown is the subdominant contribution from prompt atmospheric tau neutrinos, which is predicted at the level of 10% of the prompt muon (electron) neutrinos flux [106]. In the relevant energy range for our work, the prompt flux is strongly subdominant to the combined background from conventional atmospheric and astrophysical neutrinos.
The flux of prompt atmospheric neutrinos has been calculated independently by different authors. Two recent examples are GRRST [104] and GMVFNS [105]. The corresponding results are compared to the BERSS calculation in Fig. 1.11 and agree within $\leq 20\%$, which is smaller than the theoretical uncertainties estimated by the different author groups. Throughout this work we will use the BERSS prediction as the baseline model. Systematic uncertainties will be accounted for by allowing the flux normalization to vary (c.f. Sec. 5.2.2).

1.8 Principles of Neutrino Detection in Ice

Several different neutrino detection methods exist and target neutrinos at different energies, e.g. using scintillators [107], radio chemical methods [4], tracking calorimeters [108], Cherenkov detectors [109] and radio detectors [110]. The $\sim$ TeV-PeV energy range of interest for this work is best explored using the Cherenkov technique as done by the IceCube experiment [111]. High energy neutrinos that interact with the medium inside or nearby the detector produce charged secondary particles. These charged secondary particles during propagation produce Cherenkov light which contains information about the neutrino interaction and can be measured experimentally.

High-Energy Neutrino Interactions with Matter
At energies relevant for this work, $E \geq 100\,\mathrm{GeV}$, neutrinos interact with the detector medium (Antarctic ice, c.f. Sec. 2.2) essentially solely via *neutrino-nucleon deep inelastic scattering* (νN-DIS). In this regime, the asymptotic freedom of QCD allows for a simplified description of the nucleons in the ice (protons and neutrons bound in hydrogen and oxygen nuclei that make up water molecules), the so-called parton picture, i.e. nucleons being composed of free constituents, called partons (later identified as quarks and gluons). In this picture the neutrino interacts with the nucleon via the weak force by exchanging a $W^{\pm}$ (charged-current interaction, CC) or Z^0 (neutral-current interaction, NC) with one of the partons inside the nucleon system. The interactions are:

$$\begin{aligned} \nu_x + N &\longrightarrow x^{\pm} + \text{hadrons} \quad (CC) \\ \nu_x + N &\longrightarrow \nu_x + \text{hadrons} \quad (NC) \\ & x \in \{e,\, \mu,\, \tau\} \end{aligned} \tag{1.46}$$

[19] The difference between both results is largely due to different treatment of hadronic processes: ERS used a "colour dipole model", while BERSS uses NLO pQCD.

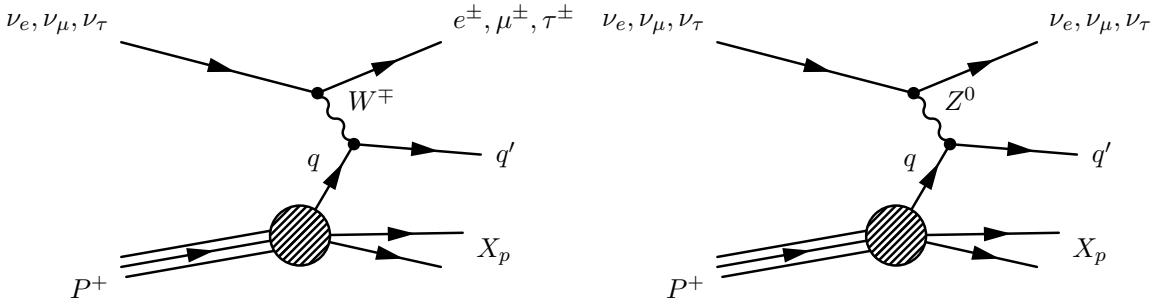


FIGURE 1.12: Feynman diagrams describing deep inelastic scattering of neutrinos off nucleons. Charged-Current interactions (left) and Neutral-Current interactions (right).

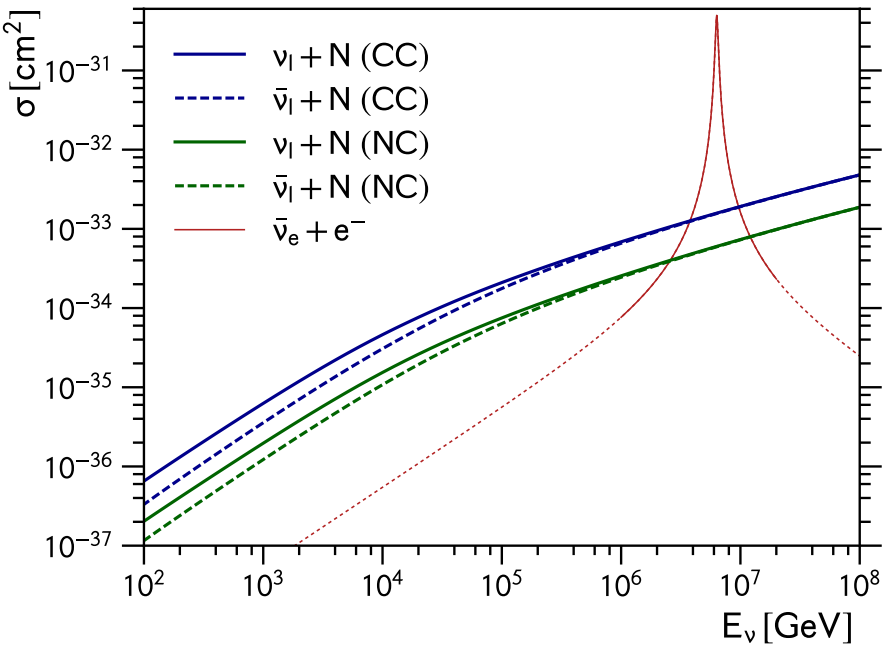


FIGURE 1.13: Neutrino-Nucleon DIS interaction cross-sections for charged-current (green) and neutral-current (blue) interactions, calculated by [112]. Resonant scattering of electron antineutrinos off atomic electrons, the Glashow Resonance [113] (red).

where we used ν_x to denote either ν_x or $\bar{\nu}_x$. Given the large (compared to the binding energy of the nucleon) energy and momentum transfer to the nucleon, the nucleon does not survive the interaction and fragments into a hadronic particle shower (hadronic cascade). Depending on whether interaction was mediated by $W^\pm$ (CC) or Z^0 (NC), the hadronic shower is accompanied by a charged leptonic or a (invisible) neutrino as final state. The corresponding Feynman diagrams are given in Fig. 1.12. The differential cross-sections can be determined from electroweak theory [114] for isoscalar targets (equal number of

protons and neutrons). For CC-interactions (inclusive) $\nu_\ell N \rightarrow \ell^\pm + X$ it reads [114]:

$$\frac{d^2\sigma}{dxdy} = \frac{2G_F^2 M E_\nu}{\pi \left(1 + Q^2/M_W^2\right)^2} \left[xq\left(x, Q^2\right) + x\bar{q}\left(x, Q^2\right)(1-y)^2\right] \tag{1.47}$$

For NC-interactions (inclusive) $\nu_\ell N \rightarrow \nu_\ell + X$ one finds [114]:

$$\frac{d^2\sigma}{dxdy} = \frac{2G_F^2 M E_\nu}{\pi \left(1 + Q^2/M_Z^2\right)^2} \left[xq^0\left(x, Q^2\right) + x\bar{q}^0\left(x, Q^2\right)(1-y)^2\right] \tag{1.48}$$

where the kinematic variables $0 \leq x \leq 1$ (Bjorken x) and $0 \leq y \leq 1$ (inelasticity) are defined as [115]

$$x = \frac{Q^2}{2ME_\nu y} \tag{1.49}$$

$$y = \frac{E_\nu - E'_{\nu,\, l^\pm}}{E_\nu} = \frac{E_{had}}{E_\nu} \tag{1.50}$$

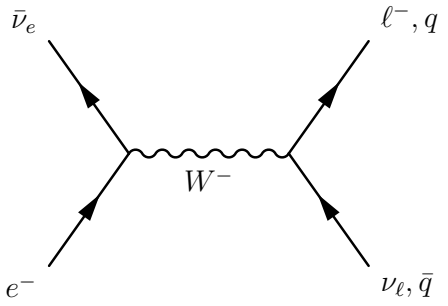
and Q^2 is defined as the negative square of the (four-)momentum transfer from neutrino to the parton. For leptonic final states l it reads

$$Q^2 = -m_l^2 + 2E_\nu\left(E_l - p_l \cos\theta_l\right) \tag{1.51}$$

Finally, the functions $q\left(x,\, Q^2\right)$ and $q^0\left(x,\, Q^2\right)$ are superpositions of parton distribution functions for the quark content in protons and are given in [114]. The former read [114]

$$\begin{aligned} q\left(x,\, Q^2\right) &= \frac{u_v\left(x,\, Q^2\right) + d_v\left(x,\, Q^2\right)}{2} + \frac{u_s\left(x,\, Q^2\right) + d_s\left(x,\, Q^2\right)}{2} + s_s\left(x,\, Q^2\right) + b_s\left(x,\, Q^2\right) \\ \bar{q}\left(x,\, Q^2\right) &= \frac{u_s\left(x,\, Q^2\right) + d_s\left(x,\, Q^2\right)}{2} + c_s\left(x,\, Q^2\right) + t_s\left(x,\, Q^2\right) \end{aligned} \tag{1.52}$$

where u, d, c and s, t, b denote the structure functions of different type of quarks, while the subscripts v and s distinguish between valence and sea contributions.
The average inelasticity $< y >$, i.e. the fraction of energy channeled into the hadronic shower, decreases with neutrino energy from ~ 0.45 (~ 0.35) at 1 TeV to ~ 0.2 (~ 0.2) in the ultra high energy limit for (anti)neutrino interactions. For this work we rely on the cross-section calculation from [112] (CSMS) based on parton distributions functions obtained with HERA [117], a lepton-proton collider at DESY (and verified against [118]). The corresponding (anti)neutrino-nucleon cross sections for CC and NC interactions are shown in Fig. 1.13 and increase with increasing neutrino energy.
While negligible almost everywhere, the total neutrino interaction cross-section is dominated near $E_\nu = 6.3\,\text{PeV}$ by the *Glashow resonance* [113], i.e. resonant scattering of electron antineutrinos of atomic electrons in the ice. The resonance energy corresponds to the rest mass of the W^- boson in the center of mass frame. The branching ratios of the W^- decay can be found in Tab. 1.14. In $\sim 70\%$ of the cases the final state does not involve a neutrino, and thus the entire energy is deposited in the detector.



W^- Decay Modes	Fraction(Γ_i/Γ_{tot})
$e^-\bar{\nu}_e$	$(10.71 \pm 0.16)\,\%$
$\mu^-\bar{\nu}_\mu$	$(10.63 \pm 0.15)\,\%$
$\tau^-\bar{\nu}_\tau$	$(11.38 \pm 0.21)\,\%$
hadrons	$(67.41 \pm 0.27)\,\%$

FIGURE 1.14: Glashow Resonance [113]: (Left) resonant s-channel scattering between electron anti-neutrino and electron, tree level. (Right) the relevant branching ratios from the W^--decay [116].

Cherenkov Radiation

Cherenkov light is emitted when charged particles propagate through a medium with refractive index $n\,(\lambda)$ at a speed $\beta \geq c_n/c_0$ that is larger than local speed of light in the medium $c_n = c_0/n$ (c_0: speed of light in vacuum). Hence one can immediately write down the threshold kinetic energy E_c above which Cherenkov emission will appear

$$E_c = E_0 \left(\left(1 - \frac{1}{n^2} \right)^{-1/2} - 1 \right) \tag{1.53}$$

For electrons in water and ice this becomes $E_c \approx 0.26\,\text{MeV}$.
The differential photon spectrum can be derived from the *Frank-Tamm formula* [119]

$$\frac{dN_\gamma}{dx\,d\lambda} = \frac{2\pi\alpha}{\lambda^2} \left(1 - \frac{1}{\beta^2 n(\lambda)^2} \right) \tag{1.54}$$

where α is the fine structure constant and x is the path length. Thus for media with refractive index that only slowly varies with the photon wavelength λ, the emission in the optical band is largest at blue, near-UV wavelengths. The angular emission profile of Cherenkov radiation is sharply peaked at the Cherenkov angle θ_c.

$$\cos\theta_c = \frac{1}{n\beta} \tag{1.55}$$

For highly relativistic particles ($\beta \approx 1$) propagating through ice ($n_{ice} \approx 1.33$ [80]) the Cherenkov angle is $\theta_c \approx 41°$. The energy loss from Cherenkov emission is completely negligible compared to other energy-loss mechanisms experienced by charged particles propagating through a medium [120]. It follows from eq. (1.54) and (1.55), that the total number of Cherenkov photons radiated in a finite wavelength band per unit track length is independent of particle energy ($\beta \approx 1$). For IceCube one finds $\sim 250\,\text{cm}^{-1}$ [121].

Neutrino Induced Cascades

The (essentially) only energy loss mechanism for electrons in matter at relevant energies above 100 GeV is energy loss via bremsstrahlung in the electric field of atomic nuclei: $Z^A + e^\pm \to Z^A + e^\pm + \gamma$. The differential cross-section for this *Bethe-Heitler process* in

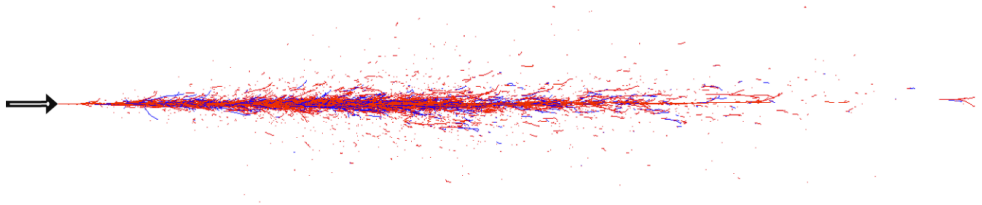


FIGURE 1.15: Simulation of an electromagnetic cascade initiated by a 100 GeV electron in ice using Geant4 (Figure from [122])

the relativistic limit $\beta \to 1$ is given by [120]

$$\frac{d\sigma}{dE_\gamma} = 2\alpha Z^2 \left(\frac{e^2}{mc^2}\right)^2 \frac{1}{E_\gamma}\left[1 + \frac{E_0'^2}{E_0^2} - \frac{2E_0'}{3E_0}\right]\left[2\ln\frac{2E_0E_0'}{mc^2E_\gamma} - 1\right] \tag{1.56}$$

with E_0 is the total energy of the incident electron. $E_0' = E_0 - E_\gamma$ denotes the total energy of the scattered electron, and E_γ is the energy of the radiated photon. Similar to the case of proton energy losses, it is useful to define the *radiation length* via $dE/ds = -E_0/X_0$, i.e. as the distance over which the energy of the electron is reduced by a factor of $1/e$. It is given approximately via

$$\frac{1}{X_0} = 4N_g\alpha Z(Z+1)\left(\frac{e^2}{mc^2}\right)^2 \ln\frac{183}{Z^{1/3}} \quad \left[\mathrm{cm}^{-1}\right] \tag{1.57}$$

where N_g is the atomic density of the material. For electrons in water one finds $\sim 36\,\mathrm{g}\cdot\mathrm{cm}^{-2}$ ($\sim 36\,\mathrm{cm}$) [120].
The emission of photons via bremsstrahlung quickly initiates an electromagnetic cascade, since the emitted photons in turn can undergo pair-production $\gamma + p \longrightarrow p + e^+e^-$ and produce pairs of electrons and positrons that will lose energy via bremsstrahlung. The interaction length of pair production is similar to that of bremsstrahlung. In the high energy limit it reads [116]

$$L_{pair} = \frac{9}{7} \times X_0 \tag{1.58}$$

The development of the electromagnetic cascade continues until the energy of the particles in the cascade drops below a certain threshold, below which energy loss is dominated by ionization.

In ice this critical energy is $E_c = 80\,\mathrm{MeV}$ [122]. For high energy cascades the total Cherenkov light emission scales linearly with the energy of the cascade, which forms the basis for energy reconstruction methods in IceCube. The longitudinal energy loss profile of a cascade in ice can, to a good approximation, be parametrized[20] by a gamma-distribution (in units of the radiation length $\xi = s/X_0$) [122]:

$$\frac{dE}{d\xi} = E_0 b \frac{(b\xi)^{a-1}\, e^{-b\xi}}{\gamma\,(a)} \tag{1.59}$$

[20] derived from GEANT4 simulations of electromagnetic cascades in ice

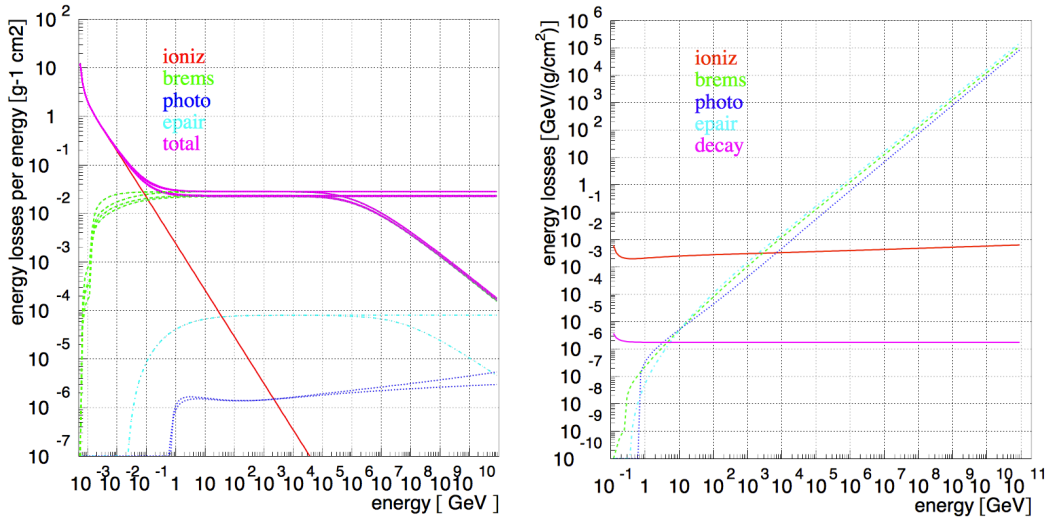


FIGURE 1.16: Energy loss rates for different energy loss mechanisms during propagation of electrons (left) and muons (right) through a medium (Figure from [123])

For water the following parameters have been found: $a = 2.0 + 1.45 \log_{10}(E_0/GeV)$ and $b = 0.63$ [122], corresponding to a fractional energy loss of 90% within a distance of $\sim 8\,\mathrm{m}$ from the vertex of the neutrino interaction for a 100 TeV cascade.

Hadronic cascades related to the fragmentation of a nucleon that was struck by the neutrino, behave qualitatively similar to their electromagnetic counterparts. The physics is more complex, since in addition to electromagnetic processes, hadronic processes (e.g. pion production and decay) contribute to the development of the cascade. Due to IceCube's coarse instrumentation the differences between purely electromagnetic showers and mixed showers of hadronic origin cannot be resolved experimentally. Hadronic cascades are thus modeled as electromagnetic cascades with reduced light yield, i.e. with lower energy to account for neutral hadrons. A light-reduction factor $E_{had}^{eff} = \eta\,(E_{had}) \times E_{had}$ has been derived from GEANT4 simulations [80]. The average light correction at an energy of $\sim 100\,\mathrm{GeV}$ ($\sim 1\,\mathrm{PeV}$) is $\eta \sim 0.7$ ($\eta \sim 0.9$).

Energy Loss of Muons in Ice

At relevant energies the energy loss of muons has a richer phenomenology than that of electrons, since different energy loss mechanisms contribute: ionization, bremsstrahlung, pair-production and photo-nuclear interactions. The first is responsible for a continuous energy loss rate close to that of minimum ionizing particles $\sim 2\,\mathrm{MeV}/(\mathrm{g/cm}^2)$ for muon energies $E\mu \ll 1\,\mathrm{TeV}$. The latter three radiative processes become important at $\sim 1\,\mathrm{TeV}$. These are stochastic processes that increase the average energy loss rate roughly linearly to the muon energy [123]

$$-\frac{dE}{dX} \approx a + bE \tag{1.60}$$

with $a = 0.249\,\mathrm{GeV/mwe}$ and $b = 0.422 \cdot 10^{-3} 1/\mathrm{mwe}$[21]. Since the muon is heavier than the electron ($m_\mu/m_e \sim 200$) the muon energy loss due to bremsstrahlung is suppressed compared to that of the muon by a factor of $(m_\mu/m_e)^2 \sim 4 \times 10^{-4}$ (c.f. eq. (1.56)). Therefore the muon appears track-like rather than shower-like in a detector like IceCube. A comparison between the different energy loss processes can be found in Fig. 1.16 for electrons (left) and muons (right).

[21] mwe ≡ meter water equivalent

Chapter 2

The IceCube Detector

2.1 Instrumentation

The IceCube detector consists of 86 cable strings that have been melted into the glacial ice at the South Pole on the Antarctic continent using a hot water drilling technique [126]. Each cable string carries 60 DOMs (Digitial Optical Modules) with a separation of 17 m and instruments the depth between 1450 m to 2450 m. The strings are distributed on a hexagonal grid with a typical separation of ~ 125 m. The resulting array of DOMs covers a total volume of about $1\,\mathrm{km}^3$ of ice, required to detect a neutrino flux at the level of the Waxman-Bahcall bound within a couple of years of operation. A schematic view of the detector geometry is given in Fig. 2.1 (top). At the bottom of the detector there exists a central region of increased instrumentation, the DeepCore infill array. It consists of eight specialized strings with reduced separation distance (72 m average) together with seven standard IceCube strings. This lowers the energy threshold from ~ 100 GeV (IceCube) to ~ 10 GeV (IceCube-DeepCore). In addition to the in-ice component, there exists an air-shower array, IceTop, on the surface. It aims at characterizing the cosmic-ray properties in the knee region of the cosmic-ray spectrum but for the purpose of this work is of no further relevance. IceCube construction started in 2005 and finished at the end of 2010. Data taking began immediately after deployment of the first string in partial detector configurations, referred to as IC-N, where N denotes the number of operational strings.

The Digital Optical Module

Each of IceCube's 5160 detection units, DOMs, consists of a Hamamatsu R7081-02 large area photomultiplier tube (PMT) with a 10 in diameter. The PMT is protected by a borosilicate glass sphere with a thickness of 1.3 cm [127] against the pressure of the ice (maximally 70 MPa) [111]. The PMT is shielded from Earth's magnetic field by a metal grid. In addition to the PMT the glass sphere houses the DOM mainboard that holds the electronics required for the data acquisition. In addition to the mainboard the glass sphere contains the flasher board the controls several LEDs within the DOM. These LEDs provide an artificial light source (405 nm) that allows for in-situ calibration of the detector. A schematic view of the DOM is given in Fig. 2.1 (bottom left).

The PMTs are operated at a nominal gain of $\gamma = 10^7$ ($U = 1.3\,\mathrm{kV}$) which produces single photon pulses of 8 mV [127]. This amplification of the single photon signal is

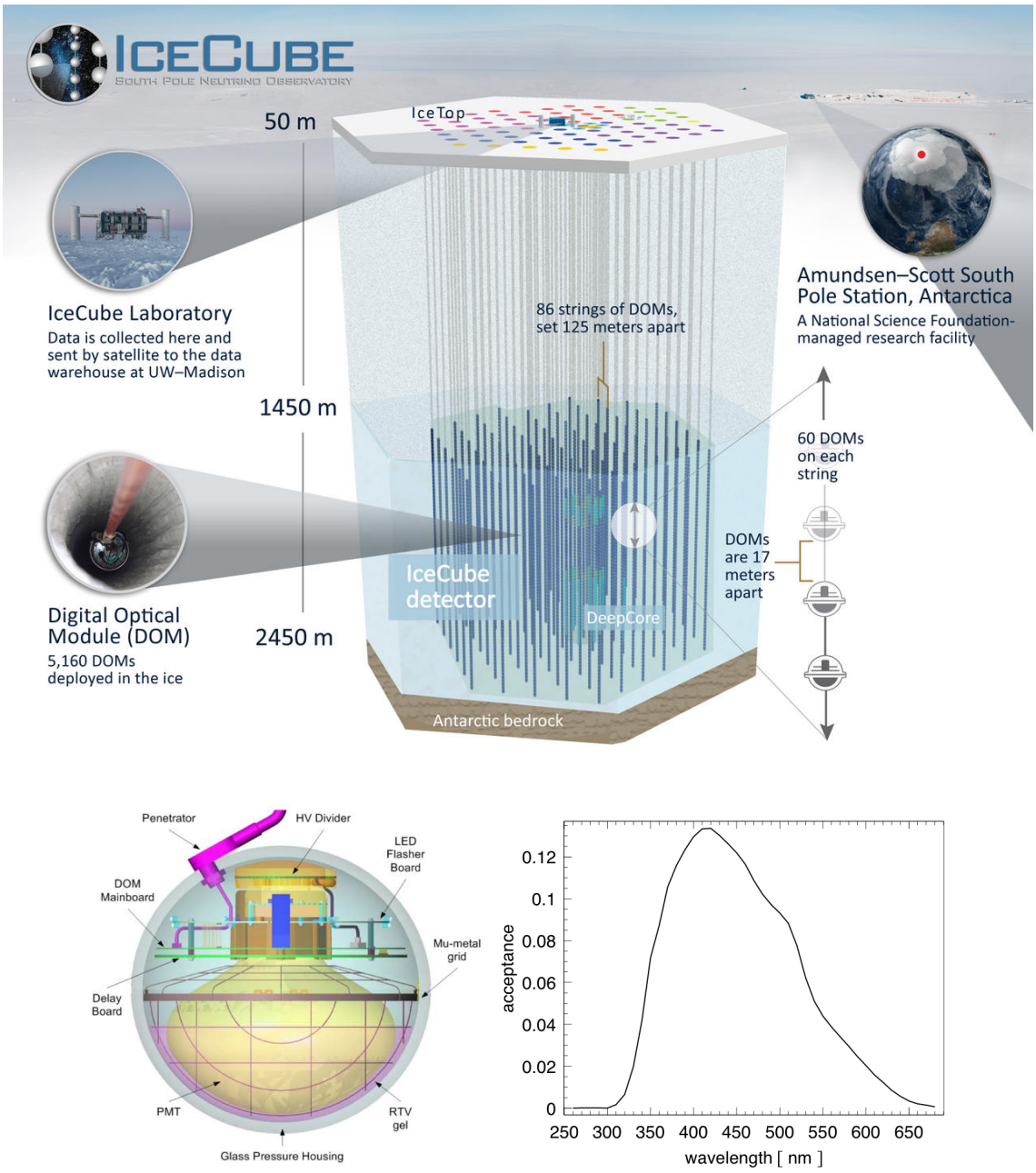


FIGURE 2.1: Overview over the instrumentation of the IceCube Observatory (top), credit: IceCube Collaboration. Structure of IceCube's Digital Optical Modules (DOMs) (bottom left, Figure from [124]. Photon Acceptance of IceCube DOM as function of photon wavelength (bottom right, Figure from [125]).

achieved in 10 dynode stages. The quantum efficiency $\eta(\lambda)$ of the IceCube PMT peaks at a wavelength of $\lambda = 390\,\text{nm}$ with $\eta_{max} = 0.25$. The (dark) noise rate of the PMTs is $\sim 300\,\text{Hz}$ at relevant temperatures ($-20°$ and $-40°$) [127] and is believed to be due to decays of radioactive elements that contaminate the glass (e.g. K^{40}). Compared to the raw PMT, the DOM has a lower quantum efficiency $\eta_{max} = 0.13$, limited by photon transmission through the glass (and optical gel that couples PMT and DOM).

The wavelength dependence of the DOM's photon acceptance is given in Fig. 2.1 (bottom right). The DOM is essentially insensitive to photons with wavelengths below (above) $\lambda \sim 300\,\text{nm}$ ($\lambda \sim 700\,\text{nm}$). The PMT response to photons has been demonstrated to be linear up to instantaneous photocurrents of 50 mA, above which the PMT enters the saturation regime and its response becomes non-linear [127]. DOMs in the saturated regime are not used in this work.

2.2 Optical Properties of the Ice at the South Pole

The glacial ice at the South Pole has been (and continuously is formed) from compacting snow on top of it. Therefore the age of the ice increases as a function of depth towards the bottom of the glacier (bed rock) at $\sim 2900\,\text{m}$. The relevant time scale for this process is $O(10^5)$y [129]. The ice is layered since the rate at which snow is deposited on top of the glacier varies with time. The compacted snow contains dust and ash impurities. Its concentration follows that of the atmosphere in the past and can be related to volcanic activity [130]. Measurements of optical properties of the ice at the South Pole at varying depths can be correlated with similar measurements at other locations on the Antarctic continent [131][132] where the age-depth relationship of the ice is known from ice-core dating [129]. The estimated age of the ice at the top of IceCube (1450 m) is $\sim 25 \cdot 10^3\,\text{y}$ and increases towards its center (1950 m) to $\sim 65 \cdot 10^3\,\text{y}$ [129]. At shallow depths below $\sim 1400\,\text{m}$ the ice contains air-bubbles that strongly scatter light. The pressure at greater depths is sufficient to compress these bubbles and gas clathrates are formed that resemble pure ice [133] and strongly increase the optical quality of the medium at IceCube depths. By using data from IceCube's LED calibration system, a model for the scattering and absorption coefficients as function of vertical depth has been developed, the *South Pole Ice* (SPICE) model [125]. The result is shown in Fig. 2.2 (magenta line, left: absorption, right: scattering). In the clearest part of the detector the measured absorption length

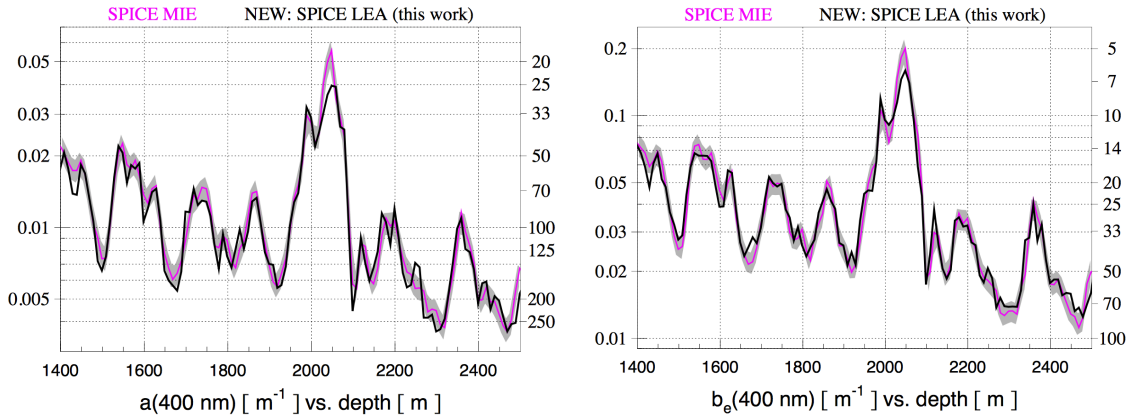


FIGURE 2.2: Optical properties of glacial ice at the South Pole: Measurements of the photon absorption and effective scattering lengths using IceCube's LED calibration system. (Figure from [128])

approaches $\sim 250\,\mathrm{m}$, which is larger than the effective scattering[1] length of $\sim 90\,\mathrm{m}$. Thus a significant fraction of light emitted by cascades is scattered before it reaches a PMT, limiting IceCube's cascade angular resolution compared to tracks. It is worth pointing out that there exists a region of very poor optical ice properties in the center of the detector with absorption (effective scattering) lengths as low as $\sim 20\,\mathrm{m}$ ($\sim 5\,\mathrm{m}$), which is called the dust-layer. Experimentally, the calibration data provides constraints at a wavelength of $\lambda \sim 400\,\mathrm{nm}$. Parameterizations exist that allow for a determination of the optical properties of the ice at different wavelengths from the ones measured at 400 nm [125]. The model includes the so-called tilt of the ice: at the bottom of the detector the ice-layers do not run perpendicular to the z-direction but are tilted. In addition, the model has been refined to account for anisotropy in the transport of light: the scattering length has been found to change with azimuthal direction presumably related to the flow of the glacier relative to the bedrock. This reduced the scattering along this direction by $\sim 8\%$ relative to the orthogonal direction [128]. We will refer to this refined model as *SPICE-Lea*, which defines the baseline for this work and is shown in Fig. 2.2 (black line, left: absorption, right: scattering) as well.

2.3 Data Acquisition, Processing and Filtering

The data recording and digitization of the PMT signals is done in a decentralized fashion [134]. The mainboard in each DOM hosts a FPGA (Field Programmable Gate Array) chip that implements the trigger logic, as shown in Fig. 2.3 (left). A simple trigger threshold corresponding to 0.25 p.e.[2] is set. If the trigger threshold is crossed, a local coincidence (LC) signal is sent to the neighboring DOMs on the same strings. If at least one neighboring DOM is triggered within a time window of 1μs, the DOM satisfies the local-coincidence criterion (HLC = hard local coincidence). If HLC is satisfied, the data is recorded over a period of 6.4μs, otherwise only a 75 ns window around the peak current within the first 400 ns is kept. Each mainboard is equipped with two ATWD (Analog Transient Waveform Digitizer) chips and one FADC (Fast Analog to Digital Converter) chip. The former can record high resolution waveforms (3.3 ns bins over $\sim 450\,\mathrm{ns}$) while the latter records the waveforms more coarsely (25 ns bins of 6.4μs). If HLC is satisfied, the first 450 ns of the event are digitized by the ATWD before the FADC takes over. AWTD digitization suffers from $29\,\mu$s dead-time and is the reason for the second ATWD chip which, together with the HLC criterion, essentially eliminates dead time. Three pre-amplifiers (multipliers 0.25, 2, 16) control the dynamic range of the ATWD.
The recorded data is sent to the surface, where the information from all DOMs is combined into events. This step depends on global trigger criteria. In this work we (implicitly) require the SMT-8 (single multiplicity trigger) criterion. These are events that contain at least 8 DOMs that recorded HLC signals within a sliding time-window of $5\,\mu\,\mathrm{s}$. Most of these triggers stem from atmospheric muons, with a rate that is more than 6 orders of magnitude larger than that of neutrinos. Since satellite bandwidth is limited ($\sim 100\,\mathrm{GB}/d$), only a fraction ($\sim 15\%$) of all triggered events can be transmitted via

[1] The effective scattering length relates to the scattering length via the average deflection angle at each scatter $\lambda_e = \lambda \frac{1}{1-<\cos\theta>}$

[2] The unit *p.e.* (# of photo electrons) is a measure of the total charge recorded by the PMT. 1 p.e. corresponds to the total charge created by one photon.

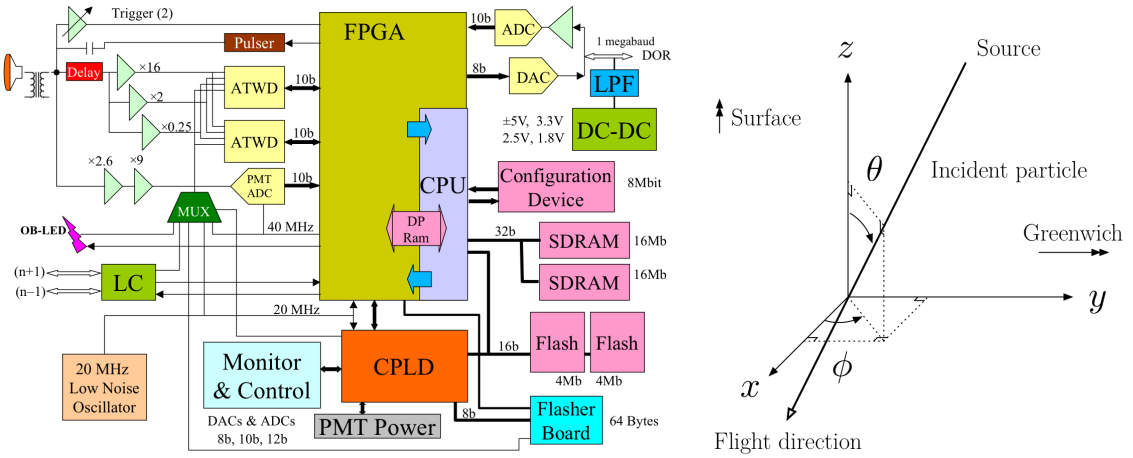


FIGURE 2.3: Schematic overview over relevant electronics inside each IceCube DOM (left, Figure from [134]). Coordinate system used to describe events in IceCube (right, credit: IceCube Collaboration).

satellite for further offline analysis in the North[3]. The decision is based on several online filters using criteria obtained from fast online reconstruction algorithms. All events studied in this work were selected by the cascade-filter ("Cascade-Level2"), designed to select events that are broadly consistent with a shower-like signature. The electron neutrino efficiency with respect to all SMT-8 filtered events is $> 70\%$ for electron neutrino events above 1 TeV and increases further with energy ($\sim 90\%$ at 1 PeV), while $\sim 98\%$ of the atmospheric muon background is removed at this step. The passing rate is $\sim 30\,\mathrm{Hz}$[4] [135]. Upon arrival at computing facilities in the Northern Hemisphere, further offline processing of the data is applied. This include additional event reconstruction and further filtering. Of interest to this work is an additional cascade-filter ("Cascade-Level3"). It retains about $\sim 40\%$ of all electron neutrinos that pass Cascade-L2 but rejects another $\sim 99\%$ of the atmospheric muon background [136].

2.4 Event Signatures and Reconstruction

The instrumentation discussed above generates two main signatures: cascades, i.e. electromagnetic or hadronic particle showers in the ice and tracks from muons propagating through the detector. Since a particle shower deposits all available energy within a small distance (several meters) compared to the distance between the strings, its Cherenkov emission appears approximately point-like. Due to photon scattering in the ice, a significant portion of the directional information related to the angular emission profile of Cherenkov radiation is lost and the photon distribution in the detector appears spherical. The remaining directional information is related to photon arrival times: PMTs in front of the cascade will on average register earlier photons (green) than the PMTs behind the cascade, which register a larger portion of late photons (blue) as shown in Fig. 2.4 (top left). The figure also shows an event view of a 2 PeV neutrino induced cascade,

[3] North = civilized world
[4] Hz=# events/s

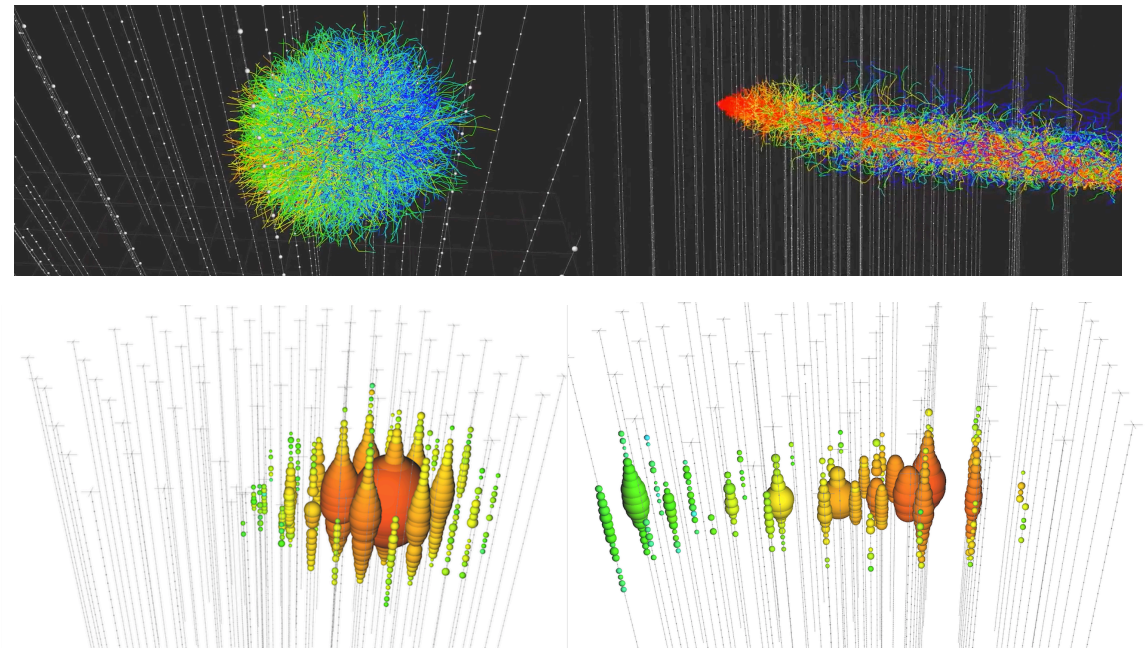


FIGURE 2.4: Event signatures in IceCube: cascades (left) and muon tracks (right). Raw photon signature from simulation (top, credit: IceCube Collaboration) and recorded signature by IceCube DOMs in data (bottom, from [5]).

recorded by IceCube (bottom left). For such events the energy can be determined to within $\sim 15\%$, dominated by systematic uncertainties. The median angular resolution, however, is rather poor ($> 10°$) [137]. Muons instead can traverse the entire detector, leaving behind a track formed by the Cherenkov photons, Fig. 2.4 (top right and bottom right). The presence of the track allows for much better directional reconstruction of tracks than cascades. For high energy through-going tracks median angular resolutions of $< 1°$ can be achieved. However muon energy estimation is more difficult since for a muon that traverses, enters or leaves the detector, only a fraction of the total muon energy loss is observed.
The different signatures allow for (limited) particle identification. Electron neutrinos produce cascades via νN DIS, and so do tau neutrinos at energies relevant for this work[5]. Muon neutrino interactions (CC) instead give rise to starting tracks: a cascade is initiated at the interaction vertex from the hadronic fragmentation of the target nucleon and a track is formed by the out-going muon - and thus these events are sometimes referred to as hybrid events. We will use both terms interchangeably. Muon neutrinos can also produce "pure" cascades via NC interactions. Finally, it is worth pointing out that down-going tracks that traverse the detector are most likely associated with atmospheric muons rather than muon neutrinos.

Waveform Calibration and Pulse Extraction

Before further analysis (reconstructions) is performed the raw data recorded in each DOM needs to be calibrated and corrected for effects from the DOM electronics. This includes baseline subtraction, i.e. the digitizer reading in absence of any signal, and gain correction

[5] different ν_τ signatures are possible at highest energies $E \gg 1\,\text{PeV}$ [138]

in order to convert the raw ADC counts into measured voltages. Additional corrections are needed to account for the transit time, the time needed for a pulse to propagate through the PMT and electronics. Finally, the waveform is corrected for transformer related distortions (droop).
The information relevant for reconstruction can be summarized in *light pulses*, characterized by the photon arrival time and total charge of the pulse, related to the number of photons. Due to finite time resolution, pulses are smeared in time $\sim 2\,\mathrm{ns}$. A bright waveform will consist of the superposition individual single photon pulses. Since the pulse shape for single photons is known from calibration, an unfolding algorithm is used to "de-convolve" the calibrated waveform in a set of individual photon pulses [137][126].

Reconstructed Variables and IceCube's Coordinate System
A cascade in IceCube is fully characterized by seven variables: vertex time t_0 (time of interaction), vertex position x_0, y_0, z_0, direction θ_0, ϕ_0 and energy E_0. Throughout this work we will measure these variables using the *IceCube coordinate system*. Its origin is centered in the detector at a depth of 1948 m, thus the top (bottom) of the detector is at $z = 500\,\mathrm{m}$ ($z = -500\,\mathrm{m}$). The angles θ_0 (zenith) and ϕ_0 (azimuth) are specified with respect to the direction where the neutrino came from, rather than the direction along which the cascade develops (opposite). Vertically down(up)-going events are associated with $\cos\theta_0 = 1.0$ ($\cos\theta_0 = -1.0$). This is shown in Fig. 2.3 (right)

Reconstruction Algorithms
The standard algorithm for the reconstruction of cascade properties, employed in this work, is *Monopod* [137][139]. Monopod relies on the method of maximum likelihood (c.f. Sec. 5.3). In particular it implements a binned poisson-likelihood function, eq. (5.15). Consider light emission from a cascade with true parameters (t_0, $\mathbf{x}_0$, θ_0, ϕ_0, E_0). The corresponding data recorded in each DOM o is segmented into bins. Within each bin, centered at time t_i, the observed number of photons (after pulse extraction) is assumed to follow a Poisson distribution with expectation $\mu_{i,o}\,(t_0,\, \mathbf{x}_0,\, \theta_0,\, \phi_0,\, E_0)$. Since the total Cherenkov light yield of a cascade is linearly proportional to its energy, the expectation $\mu_{i,o}$ can be related to that of a template cascade $\Lambda_{i,o}\,(t_0,\, \mathbf{x}_0,\, \theta_0,\, \phi_0)$ with energy 1 GeV

$$\mu_{i,o}\,(t_0,\, \mathbf{x}_0,\, \theta_0,\, \phi_0,\, E_0) = E_0 \times \Lambda_{i,o}\,(t_0,\, \mathbf{x}_0,\, \theta_0,\, \phi_0) + \rho_o \tag{2.1}$$

where we added the noise contribution ρ_o to bin i of DOM o.
A large scale simulation of template cascades with varying (discrete) directions and positions in the detector has been performed using *Photonics* [140] and the corresponding light yield tabulated as a function of true cascade parameters. The resulting tables were interpolated using B-spline surfaces [137]. This allows to obtain for each DOM o and time bin i the expected light yield $\Lambda_{i,o}\,(t_0,\, \mathbf{x}_0,\, \theta_0,\, \phi_0)$. The log-likelihood function (up to terms independent of the parameters) becomes

$$\log L\,(t,\, \mathbf{x},\, \theta,\, \phi,\, E\,|\,\mathbf{X}) = \sum_{o}^{\text{DOMs}} \sum_{i}^{\text{bins}} q_{i,p} \log\left(E \times \Lambda_{i,o} + \rho_{i,o}\right) - \left(E \times \Lambda_{i,o} + \rho_{i,o}\right) \tag{2.2}$$

where $\mathbf{X}$ denotes the measured pulses (total charge and time) in all DOMs. Minimizing the negative log-likelihood numerically using gradient descent methods (in combination

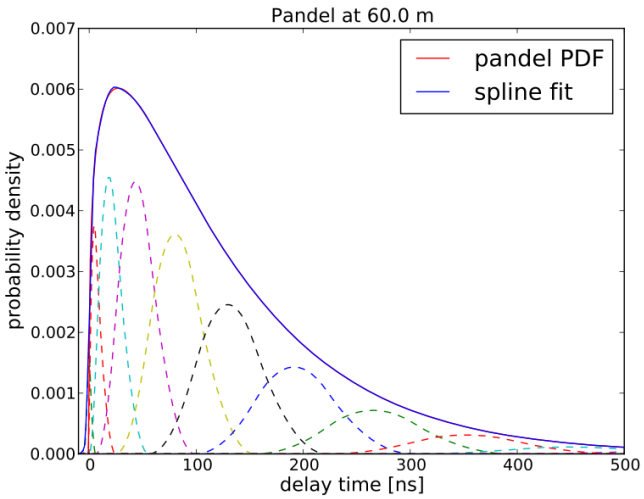


FIGURE 2.5: Analytic photon delay time distribution (Pandel [143]) with non-parametric spline fit. (Figure from [144])

with Newton's method to solve for the energy in each step) yields the best-fit, reconstructed values $\hat{t}$, $\hat{\mathbf{x}}$, $\hat{\theta}$, $\hat{\phi}$, $\hat{E}$.

Several other reconstruction algorithms exist that are faster but less precise. Given their speed, they can be applied to large data samples. For the purpose of this work they are used to either define cut variables, or more importantly, to provide a starting point for Monopod. The performance of numerical minimization in several dimensions and complicated likelihood spaces hinges strongly on the accuracy of the starting point.
A simple estimate of a cascade vertex can be obtained from the *center-of-gravity* (cog) calculation. It is a simple weighted average of the three-dimensional positions of all DOMs that recorded light during the event.

$$\mathbf{x}_{cog} = \frac{\sum_o w_o \mathbf{x}_o}{\sum_o w_o}, \quad w_o \equiv \left(\sum_o Q_o\right)^{\alpha} \tag{2.3}$$

where Q_o is the total charge recorded by DOM o. Assuming spherical emission of light at the local speed of light, the vertex time t_0 corresponding to the cog vertex estimate can be obtained [141].
A first guess for the cascade direction can be obtained from considering the *tensor-of-inertia* of the DOM hits, where the total charge Q_o, observed in DOM o, serves as mass.

$$I_{ij} = \sum_o Q_o^{\alpha} \left(\delta_{ij}\mathbf{x}_o^2 - x_{o,i}x_{o,j}\right) \tag{2.4}$$

An improved estimate of the interaction vertex $\mathbf{x_0}$ and time t_0 can be obtained from *CascadeLlhVertexFit* [142]. It uses the arrival time of the first pulse t_{o1} observed in each DOM to formulate the following negative log-likelihood function

$$-\ln L\left(t_0, \mathbf{x_0} | t_{o1}, \mathbf{x}_o\right) = -\sum_o \ln p_o\left(t_{o1}, \mathbf{x}_o | t_0, \mathbf{x_0}\right) \tag{2.5}$$

where the photon arrival time distribution is given by an analytic function, the *Pandel*-pdf [143], which assumes a homogenous medium and ignores the complexity of the realistic ice-model. In the calculation only time delays with respect to the naive arrival time, expected from direct propagation at the local speed of light, are considered. An example is shown in Fig. 2.5.

In the context of this work, muon tracks are considered as background. Correct reconstruction of muon tracks helps identifying them. Therefore in addition to the cascade reconstruction methods discussed above, we also incorporate track reconstruction algorithms into this analysis. From the perspective of reconstruction, the difference between cascades and tracks is that the latter serves as a light source that moves through the detector at the speed of light, while the former is considered a stationary source.
A simple track reconstruction algorithm is the *LineFit* [145]. It considers a planar light front that propagates through the detector with velocity $\mathbf{v_0}$ along the track direction. Thus the physics of Cherenkov emission, photon transport in inhomogenous ice and detector noise is ignored. The problem then reduces to linear regression in three dimensions, with time as predictor.

$$\hat{\mathbf{x}}_\mathbf{0}, \hat{\mathbf{v}}_\mathbf{0} = \underset{\mathbf{x},\mathbf{v}}{\operatorname{argmin}} \left\{ -\sum_i \left(\mathbf{x}_i - (\mathbf{x} + \mathbf{v} \cdot t_i)\right)^2 \right\} \tag{2.6}$$

where the sum runs over all hits in all DOMs at position $\mathbf{x}_i$ and time t_i.
The result from LineFit serves as seed for a likelihood based track reconstruction: *SPEFit* [146]. The implementation is very similar to that of CascadeLlhVertexFit in that the likelihood uses the analytical Pandel delaytime pdf. The difference between both is simply that SPEFit likelihood considers time delays with respect to a moving source, while the cascade reconstruction considers a stationary light source [80]. Throughout this work we do not use track reconstruction algorithms that attempt to reconstructed the muon energy. The final sample will be dominated by cascades and the final energy measurement is performed using the Monopod cascade reconstruction.

2.5 Event Simulation

Data analysis in IceCube relies on extensive Monte Carlo simulations that model all relevant processes involved in the generation of experimental data:

1. Generation of Primary Particle, Propagation and Interaction
2. Charged Lepton Propagation
3. Propagation of Cherenkov Photons
4. Response of Detector Electronics to Light Signals

The simulation of neutrino events is performed using the *neutrino generator* (NuGen) software. NuGen is a customized version of the ANIS neutrino generator [147], designed

for better compatibility with other IceCube software. NuGen generates neutrinos within some energy range following a probability distribution determined from the assumed neutrino flux, usually a power-law. Each neutrino is injected at Earth's surface and then propagated towards the detector taking into account all standard model neutrino interactions in matter. Here the Earth is described by the Preliminary Earth Model (PREM). Upon arrival at the detector a neutrino interaction is simulated with a random vertex position a long the neutrino trajectory crossing the detector and secondary particles are generated.
Muons produced inside (or near) the detector are then propagated and their energy losses simulated using the MMC [123] or Proposal [148] packages. Cascades below 1 TeV are simulated as stationary light emitters, while the elongation of cascades at higher energies is taken into account by placing discrete energy losses (stationary emitters) along the trajectory following the longitudinal shower profile from eq. (1.59).
The computationally most expensive step is the simulation of the photon transport through the inhomogeneous ice, since each individual photon has to be tracked. This is performed using the software packages CLsim [149] or PPC [150].
Finally the response of the detector electronics to the Cherenkov photons is simulated. This includes simulation of detector noise during the time period of the event, physical processes related to the PMT and the digitization process (DOM electronics). The resulting simulation data is then treated like real data and subject to further filtering and processing.

Background events (atmospheric muons) are generated using CORSIKA (Cosmic Ray Simulations for Kascade) [151]. CORSIKA simulates particle production in cosmic ray induced atmospheric air-showers. The composition is taken into account by focusing on five representative cosmic ray primary nuclei (H, He, N, Al, Fe), each representing the corresponding mass group (c.f. Sec. 1.1). Of particular interest is the muon yield (muon bundles and single muons) from such showers. Once the muons reached the surface, they are further propagated towards the detector using the MMC/Proposal packages.
Since full simulation of the air-shower development is computationally expensive, resources can be saved by generating muon events from parameterizations of the muon yield near the detector derived from CORSIKA and Proposal. Such parameterizations have been obtained in [139] and implemented in MuonGun, a fast generator of single muon events. However, since MuonGun does not account for any contribution from bundles of muons, it can not describe experimental data - unless filter conditions (event selection) can be shown to efficiently remove muon bundles.

Chapter 3

Event Selection

The dataset analyzed in this dissertation consists of IceCube data recorded in a four year period covering the **years 2012 to 2015**. The selection criteria presented in this chapter have been developed using only a fraction (10%) of the total dataset in order to avoid introducing selection bias by the analyzer. This so-called blind analysis approach is a standard requirement within the IceCube Collaboration[1]. Our goal is to devise filtering criteria that reduce the atmospheric muon background from being dominant over the neutrino component (factor of $\sim$ 450) to a subdominant level ($<$ 10%). Of particular interest are neutrino-induced cascades (c.f. Sec. 2.4), since they have a lower contribution from atmospheric neutrinos compared to muon neutrino induced (starting) tracks (c.f. Sec. 1.7) and provide excellent sensitivity to the astrophysical flux of electron and tau neutrinos. In addition to a cascade-signal sample (c.f. Sec. 3.2.7 and Sec. 3.2.8), we will define two control samples that constrain the properties of the non-electron (and non-tau) neutrino backgrounds. First, muon neutrinos that interact via the neutral current channel (NC) produce cascades that give rise to an irreducible background from atmospheric neutrinos[2]. By isolating a control sample of starting tracks (c.f. Sec. 3.2.6) consisting almost entirely of muon neutrinos from conventional atmospheric origin interacting via the charged-current channel, we will be able to effectively constrain the corresponding muon neutrino NC contribution to the cascade signal sample, since the ratio between the DIS interaction cross-sections for NC and NC interactions are known well [112]. Finally, in order to characterize and constrain the remaining residual contribution from atmospheric muons to the final cascade sample, we will define a muon control sample which consists mostly of single muon events with a cascade-like event signature (c.f. Sec. 3.2.8).

The biggest challenge throughout the development of this event selection was posed by the comparably small amount of available cosmic-ray air-shower simulation (CORSIKA [151]). We were able to overcome this problem by demonstrating that, as a function of increasing neutrino purity, the cosmic-ray background reduces to single muons. This selection identifies bundles of muons as background which we remove from the sample (c.f. Sec. 3.2.4). This allows to replace full air-shower simulation by a faster and more efficient single muon simulation (MuonGun [139]).

[1] There are variations as to what defines a blind analysis across different experiments and different fields.

[2] This argument ignores the suppression from the atmospheric neutrino self-veto effect, c.f. Sec. 4.2.3

Another difficulty is created by improvements in the IceCube online and offline filter and processing streams that introduce variations in the filter rates across the early years (2011-2013) and impact the very background-like region of this analysis. We will show that this event selection unifies the rates observed in data (and simulation) as well as the shapes of all relevant observable distributions for all years (c.f. Sec. 3.2.4). This allows to perform the final statistical analysis using one set of simulations for the entire data sample which significantly lowers the computational burden.
The central piece of the event selection is a multi-class gradient-boosted decision tree model [152] that we apply to each event satisfying a set of pre-selection criteria (c.f. Sec. 3.1). Based on this model, for each event we assign three scores (class membership probabilities) corresponding to the three different classes of events: cascades, starting tracks and atmospheric muons (c.f. Sec. 3.2). Using these scores we either remove the event from the analysis, or we assign them into one of the three mutually exclusive samples: the cascade signal sample (c.f. Sec. 3.2.7) and the two control samples: one for starting tracks (c.f. 3.2.6) and one for single muons with a nearly cascade-like signature (c.f. Sec. 3.2.8).
Finally, we will show that another event selection, developed at Stony Brook University [153], specifically designed to target high energy cascades with $E > 60\,\text{TeV}$, provides higher signal-efficiency than the one presented here at highest energies. Thus for the measurement of the astrophysical neutrino flux (c.f. Sec. 6) will use both, the low-energy selection developed here and the dedicated high energy selection from [153] with a transition at a reconstructed energy $E_{rec} = 60\,\text{TeV}$ (c.f Sec. 3.3).

3.1 Pre-Selection: Cleaning the Data Sample

IceCube's online and offline filtering streams, while significantly reducing the muon background rates compared to the rate at which the detector is triggered, can be viewed as rather loose selections. They are not designed to provide a set of high quality neutrino events, rather, they are intended to be a common starting point with reduced data volume for the different analyses that target widely different physics topics and disparate (by orders of magnitude) energy ranges. Here we define high-quality events as events that are well-reconstructed and well-described by simulations. As we will show, the agreement between data and MC simulations at the starting point of this analysis, IceCube's cascade filter, is rather poor in some regions of observable space, for example mis-reconstructed events at high energies, dominated by mis-reconstructed muon bundles. Thus at this stage we cannot consider the MC simulations to be a sufficiently accurate representation of all important data-generating processes involved. Our intention is to model the differences in observable space between the different types of events (signal and background) by means of machine learning. Thus, before fitting such a model, we first need to improve upon the agreement between the observed data and our simulations. This is done by identifying and removing events that appear to cause the observed discrepancies. These cleaning cuts are discussed below and we will subsequently refer to them as "Level4"-cuts.
Our approach to cleaning cuts differs from other IceCube analyses in that we have to satisfy one important condition. We have to find cleaning cuts that do not reject too much background (and thus do not drastically lower the overall data rate). This might appear contradictory at first, since rejecting background is what we want after all. The

reason is that fitting a complex machine learning model requires a sufficiently large sample of simulated background events. Rejecting too much background at the cleaning stage of this analysis would decrease further the already small sample of available air-shower simulation (CORSIKA). Our pre-selection reduces the background rate (data rate) by $\sim 30\%$, while the final machine learning selection achieves a reduction of background by ~ 4 orders of magnitude.
Finally, during the development of the event selection, essentially no CORSIKA simulation was available matching the exact detector/filter configurations for the years 2012-2015, that are relevant for this work. The most recent and sufficiently large simulation dataset uses the detector configuration from the year 2011, and hence, this dataset is/was used for the development of all selection criteria. The agreement between data/MC in the background region is then evaluated using the corresponding 10% test-sample of the 2011 experimental data. As we will show using newer simulations that were produced during development of this analysis, the data/mc agreement obtained after the cleaning cuts generalizes well to the 2012-2015 detector configurations.

3.1.1 Definition of Variables

Throughout this chapter a set of variables/observables will be used that can be grouped according to their purpose:

Group 1: Estimated/Reconstructed Energy Deposition

- **Monopod.energy** (from Monopod **cascade** reconstruction)
- **Total Charge of all PMTs**

Group 2: Estimated/Reconstructed Position of Interaction Vertex

- **Monopod.z** (from Monopod **cascade** reconstruction)
- **Monopod.XYScale** (based on Monopod **cascade** reconstruction)
- **Depth of First Photon Hit**
- **Depth of DOM with Maximum Charge**

Group 3: Reconstructed Arrival Direction (Zenith Angle)

- **Monopod.zenith** (from Monopod **cascade** reconstruction)
- **LineFit.zenith** (from LineFit **track** reconstruction)
- **Track.zenith** (from SPEFit **track** reconstruction)

Group 4: Quality of Reconstruction

- **Cascade.rlogl** (from CascadeLlhVertexFit **cascade** reconstruction)
- **Track.rlogl** (from SPEFit **track** reconstruction)
- **Reconstructed Vertex Separation** (from CascadeLlhVertexFit and Monopod **cascade** reconstructions)

Group 5: Identification of Incoming/Outgoing Muons

- **Photon Delay Time** (based on Monopod **cascade** reconstructed vertex position)
- **Veto Track Charge** (based on Monopod **cascade** reconstructed vertex position)
- **Starting Track Charge** (based on Monopod **cascade** reconstructed vertex position)

These variables represent a mixture of rrobust low-level estimates of event properties with more elaborate and precise high-level calculations. Similarly this set includes variables that are more appropriate for the reconstruction of cascades, the signature of interest for this work, and variables that are better suited for track reconstruction that help to identify, characterize and reject muon background. All variables, except for the energy estimate from Monopod (Monopod.energy) and the DOM with maximum charge (MaxChargeDOM) will be used in the gradient boosted decision tree model (c.f. Sec. 3.2). Each variable is explained below.

Monopod.energy, Z-position, XYscale and Zenith

Monopod (eq. (2.1)) is the most sophisticated maximum likelihood reconstruction technique employed in this work. Its reconstructed energy, reconstructed z-coordinate of the vertex position and reconstructed zenith angle are used directly as observables. As we will show later on, atmospheric muon background contributes strongest near the detector boundaries and for down-going trajectories. However for the purpose of distinguishing neutrinos from atmospheric muons not all information from the reconstructed vertex position (x,y-coordinates) is useful. As far as the x,y-coordinates are concerned, almost all information is contained in the distance of the reconstructed vertex position in the XY-plane to the detector boundary. Thus it is possible to reduce the dimensionality by one. Following [154] we introduce the variable "XYscale" which measures the position of the Monopod reconstructed vertex in the XY-plane relative to the boundary on a linear (dimensionless) scale bounded from below at XYscale = 0 (vertex located in the center of detector). The polygon that defines the detector boundary corresponds to XYscale = 1, i.e. events reconstructed with vertices corresponding $0 \leq \text{XYscale} \leq 1$ are considered inside the detector while XYscale > 1 are considered outside the detector (in the XY-plane). Additional information about the details of the implementation can be found in [154].

Total PMT Charge: Q_{tot}^{HLC}

The total charge per event collected by all IceCube DOMs constitutes a simple variable that correlates with the energy deposited by charged particles in the detector: low-energy events produce less charge than high-energy events. The total charge does not allow for a precise estimate of the energy deposit because it also depends on the position of the interaction relative to the nearest DOMs. Events close to a string produce more visible charge in the detector than events of similar energy with vertices at larger distances to the nearest strings. While dedicated energy reconstruction algorithms provide far superior energy estimates, the total charge is a good indicator as to the information content in the event. Brighter events can be reconstructed with better accuracy. Finally, when calculating Q_{tot}^{HLC} we only consider those pulses from the PMTs that satisfy a hard local coincidence (HLC) criterion (c.f. Sec. 2.3) to reduce the contribution from detector noise. The total charge is measured in units of number of photo electrons (p.e.) equivalent the total charge.

DOM with Maximum Charge

IceCube DOMs are enumerated along each string by integers ranging from 1 (top of detector) to 60 (bottom of detector). The variable "VetoMaxChargeDOM" refers to the DOM number of the DOM that collected the largest amount of charges compared to all other DOMs. For cascade-like energy deposits one expects this to be the DOM that is closest to the interaction vertex with some dependence as to the direction of the photon emission, since DOMs behind the Cherenkov cone receive less charge. Thus, this variable is rough estimate of the z-position of the event inside the detector. In particular it does not require assumptions as to the nature of the energy deposit: cascade-like or track-like.

Depth of First Hit

The z-position of the DOM that registers the first photon hit in the event is very informative about whether a cascade event inside the detector might be associated with a incoming track. The closer the depth of the first photon hit is to the top of the detector, the more likely the event is cause by a muon that penetrated into the detector from above. Since individual photon hits could also be due to detector noise, it is necessary to first apply a noise cleaning algorithm to all registered pulses from all DOMs in the event that removes noise pulses. We will subsequently refer to this variable as "VetoMaxChargeDOM".

LineFit Zenith Angle

The LineFit (c.f. eq. (2.6)) provides a very simple estimate for the arrival direction for track-like events. In the context of this work, we only utilize the reconstructed zenith angle. Atmospheric muons are more likely to be associated with down-going directions

than neutrino events. LineFit does not provide a good directional estimate for cascade-like events and thus its distribution appears more uniform. Therefore the signal-to-background ratio depends on the zenith angle reconstructed by LineFit.

SPEFit Zenith Angle

Another, more elaborate estimate of the arrival direction of track-like events can be obtained from the SPEFit [146]. The scope of this variable for the purpose of this work is similar to the one of the LineFit zenith angle. While both variables are correlated for well-reconstructed tracks, their estimates differ for events that are less well described by a track hypothesis. Thus the combination of both variables carries extra information beyond the individual directional estimates.

Track rlogl (SPEFit rlogl)

The SPEFit also provides an estimate of its fit quality: the SPEFit reduced log-likelihood. Smaller values are indicative of better agreement between the track hypothesis and the data than larger values. The number is calculated from the absolute value of the SPEFit log-likelihood function at its minimum but normalized to the total number of DOMs that registered photon hits.

Cascade rlogl (CascadeLlhVertexFit rlogl)

This variable is similar in interpretation to the one above but assumes a cascade hypothesis and thus measures the quality of the cascade fit. In particular the likelihood function of the CascadeLLhVertexFit reconstruction algorithm (eq. (2.5)) is used.

Photon Delay Time

The Photon Delay Time calculation is based on the reconstructed vertex position $\boldsymbol{x_0}$ (and time t_0) from the Monopod cascade reconstruction algorithm. In this calculation a rudimentary cascade model is assumed: isotropic, point-like photon emission at time t_0 and vertex position $\boldsymbol{x_0}$. Let $\boldsymbol{x_i^{DOM}}$ be the position of DOM i in the detector array and t_i^{DOM} be the first hit recorded in DOM i. Assuming photon propagation at the speed of light in the medium c_{ice} one can define the time of direct travel from $\boldsymbol{x_0}$ to $\boldsymbol{x_i^{DOM}}$: $t_i^{geo} = d/c_{ice}$ where $d = d(\boldsymbol{x_0}, \boldsymbol{x_i^{DOM}})$ is the distance between vertex position $\boldsymbol{x_0}$ and DOM position $\boldsymbol{x_i^{DOM}}$. The (minimum) delay time Δt_{ice}^{min} is then given by

$$\Delta t_{ice}^{min} = \min_{i \in \text{all DOMs}} \left(t_i^{DOM} - t_i^{geo} \right) \tag{3.1}$$

Under ideal conditions (e.g. the true vertex is perfectly known, the light emission is instantaneous and point-like etc.) one expects $\Delta t_{ice}^{min} \geq 0$ns, because under these assumptions photons can not arrive at the DOMs before the geometric time (the local speed of light

needs to be obeyed). They can, however, arrive later due to scattering during propagation from source to receiver. There are four processes that can lead to $\Delta t_{ice}^{min} < 0$ns. First, noise hits obviously are not covered by the considerations above. They are effectively suppressed by requiring a hard local coincidence (HLC) criterion (c.f. Sec. 2.3) and additional pulse cleaning. Second the necessarily imperfect vertex reconstruction (the true interaction vertex position is unknown) can overestimate $d = d(\boldsymbol{x_0}, \boldsymbol{x_i^{DOM}})$ even if the event under consideration is a perfect cascade. The third case is due to muon neutrino charged-current interactions. They produce a charged muon that is not bound by the local speed of light in the medium c_{ice} and thus can out-run the the light front and lead to the detection of Cherenkov photons in DOM i at $t < t_i^{geo}$. Finally the most interesting contribution is case four: atmospheric muon background can create early photon hits (mostly near the detector boundaries) before being reconstructed somewhere inside the detector. Thus this variable provides good discrimination against atmospheric muon tracks.

Veto Track Charge

Atmospheric muons penetrate the detector from the outside with down-going trajectories. We will show in Sec. 3.2.4 that the most difficult, i.e. cascade-like, muon background stems from atmospheric muons with a dim track and large stochastic energy loss within the detector. A good estimate of the event's interaction vertex can be obtained from Monopod. Under the assumption that the event is associated with an atmospheric muon, its track would have to go through the vertex of the energy deposition. To fully define that track its direction is needed. If we were to know the true trajectory of the hypothetical muon, one could check for (rare) photon detections in the DOMs along the track that are causally connected. Assuming a dim track, the standard track reconstruction algorithms are unfortunately very likely not to find the correct direction. Following [139], we therefore search for photon detections that are possibly related to one out of 104 different down-going directions. The "Veto Track Charge" is then defined as the total collected PMT charge along the track with the largest associated collected charge in comparison to all other 103 tested trajectories. For neutrinos this charge is expected to be smaller[3] than for muons, since only detector noise contributes. More details about the implementation can be found in [139].

Starting Track Charge

The "Starting Track Charge" allows for separation of starting tracks (charged current muon neutrino interactions) from cascades. The idea is similar to the one of "Veto Track Charge". While the latter searches for photons consistent with an incoming, down-going muon track, the former identifies photons consistent with an outgoing track, testing all directions. Photon hits along a possible muon trajectory are counted as possibly associated with the track if are causally consistent with the muon track, but inconsistent with the light front from the cascade. The distinction is possible because muons in ice

[3] It is smallest for astrophysical neutrinos since they are not accompanied by muons from atmospheric airshowers, while atmospheric neutrinos might be (c.f. Sec. 4.2.3 "Atmospheric Self-Veto Effect").

propagate faster than the local speed of light in the medium (also exploited in the Photon Delay Time variable). More details about the implementation can be found in [139].

Reconstructed Vertex Separation

As discussed in Sec. 2.4 the reconstruction of cascade properties is done in stages. The reconstruction results obtained with simple algorithms are used as a starting point (seed) for more complex calculations. For true cascade events on expects the latter to be a refinement of the former, but not to be significantly different. If the event is not a cascade, but a track mis-reconstructions are expected and stronger discrepancies between the different reconstructions will be observed. Here we introduce the separation distance in the XY-plane between the vertices as reconstructed by CascadeLlhVertexFit and Monopod as measure of discrepancy.

All of these variables (except for the last) have been used in previous IceCube analyses and found to perform well [80][154][139][155].

3.1.2 Level 3: The Cascade Filter Stream

The very first selection criterion defines the starting point of this analysis. We require all events to pass the designated, collaboration-wide online and offline cascade filters (c.f. Sec. 2.3). In particular we will only analyze events that are retained in the **contained**, **single** (non-coincident) **event branch** of the offline **cascade-Level3 filter**. All other events are discarded right away. Fig. 3.1 shows the reconstructed (Monopod) energy distribution for the 10% test-samples of the years 2011 (top left) and 2012 (top right) in comparison to the corresponding air-shower simulation prediction from CORSIKA. One immediately notices an excess of data events over the expectation from CORSIKA that becomes stronger with increasing energy and is most noticeable for reconstructed energies above 10 TeV. Fig. 3.1 also shows the total collected charge by all PMTs, Q_{tot}^{HLC} (bottom), which does not show similar behavior. This can be interpreted as a first hint that the excess events are due to mis-reconstructions, which appear to differ between data and simulation. We intend to suppress these events with the subsequent cleaning cuts. It is worth pointing out the good shape agreement for the total charge between data observation and simulation prediction. IceCube has recently re-calibrated the single photon charge-response function (c.f. Sec. 2.1) individually for all PMTs and is currently in the process of re-processing all experimental data. These new calibrations are already taken into account in the data processing and filtering scheme developed for this analysis. The total passing rates for experimental data (all years + 2011) and CORSIKA prediction (2012 + 2011) are given in Tab. 3.1 and match within $< 15\%$. At this stage of the event selection neutrinos are negligible. They contribute about 1 in 500 events. The neutrino selection efficiency of all Level-4 pre-cuts will be discussed in Sec. 3.1.7.

Rate in Year	**2015** [mHz]	**2014** [mHz]	**2013** [mHz]	**2012** [mHz]	2011 [mHz]
Data (10%)	$\mathbf{154.5 \pm 0.2}$	$\mathbf{155.2 \pm 0.2}$	$\mathbf{154.7 \pm 0.2}$	$\mathbf{193.9 \pm 0.3}$	146.7 ± 0.2
CORSIKA	n/a	n/a	n/a	$\mathbf{183.5 \pm 0.3}$	129.1 ± 0.5

TABLE 3.1: After Cascade-Level3 (singles, contained branch) selection: Passing rates of experimental data for data taking periods 2011-2015 in comparison to estimated passing rates from Monte Carlo simulations (CORSIKA). n/a ≡ simulation not available.

3.1.3 Level 4A: Treating the Dust Layer

The so-called dust-layer (c.f. Sec. 2.2), a region of ice contaminated with dust particles that reduce its optical quality, is particularly difficult to model with regard to atmospheric background. Due to the reduced absorption length at depths between $\sim$ 2000 m and $\sim$ 2200 m reconstruction algorithms tend to fail for most events. In particular incoming muon (bundles) with faint tracks are difficult to identify since they leave little trace in that region. Faced with possibly mis-reconstructed vertices we use the DOM with the maximum charge to identify events that deposit energy in the dust layer. The distribution of this VetoMaxChargeOM variable is shown in Fig. 3.2 (top left). The DOMs 31-38 are located in the dust-layer and a clear excess of the experimental data over the

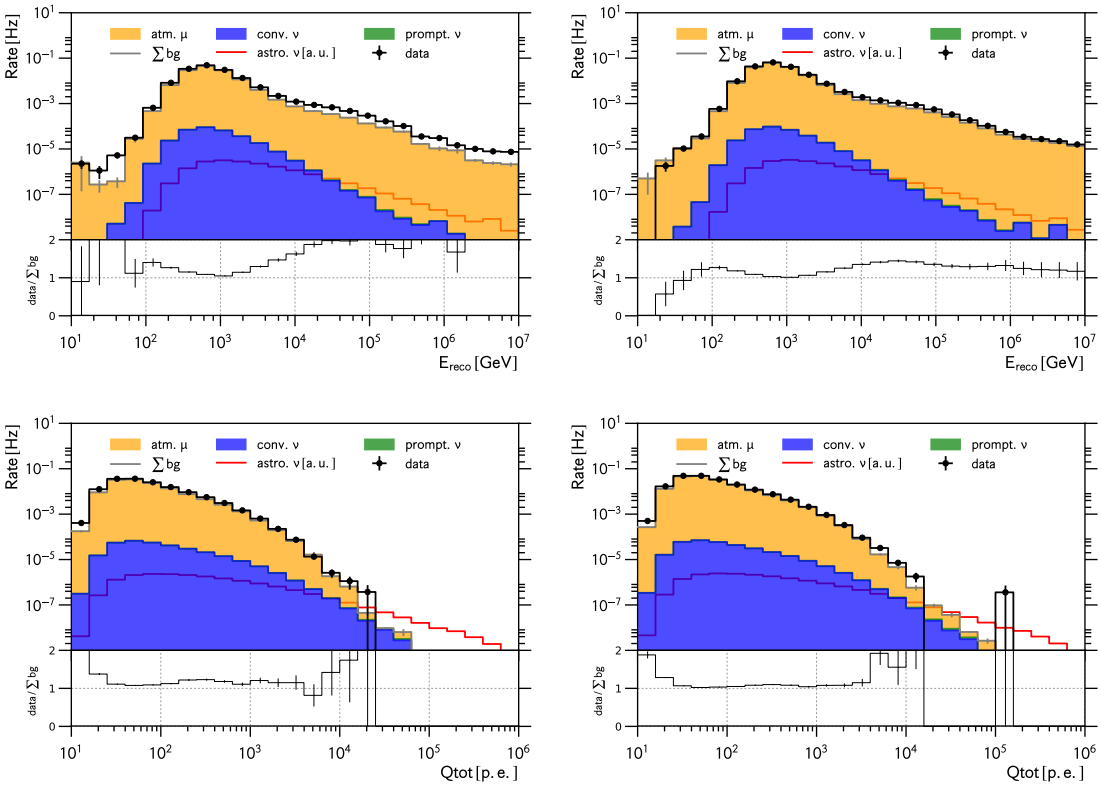


FIGURE 3.1: After Cascade-Level3 (singles, contained branch) selection: Observed distributions of reconstructed energy (top) and total PMT charge (bottom) in data period 2011 (left) and 2012 (right) using 10% of available data statistics. 2011 (left) and 2012 (right) Monte Carlo simulations: CORSIKA and NuGen.

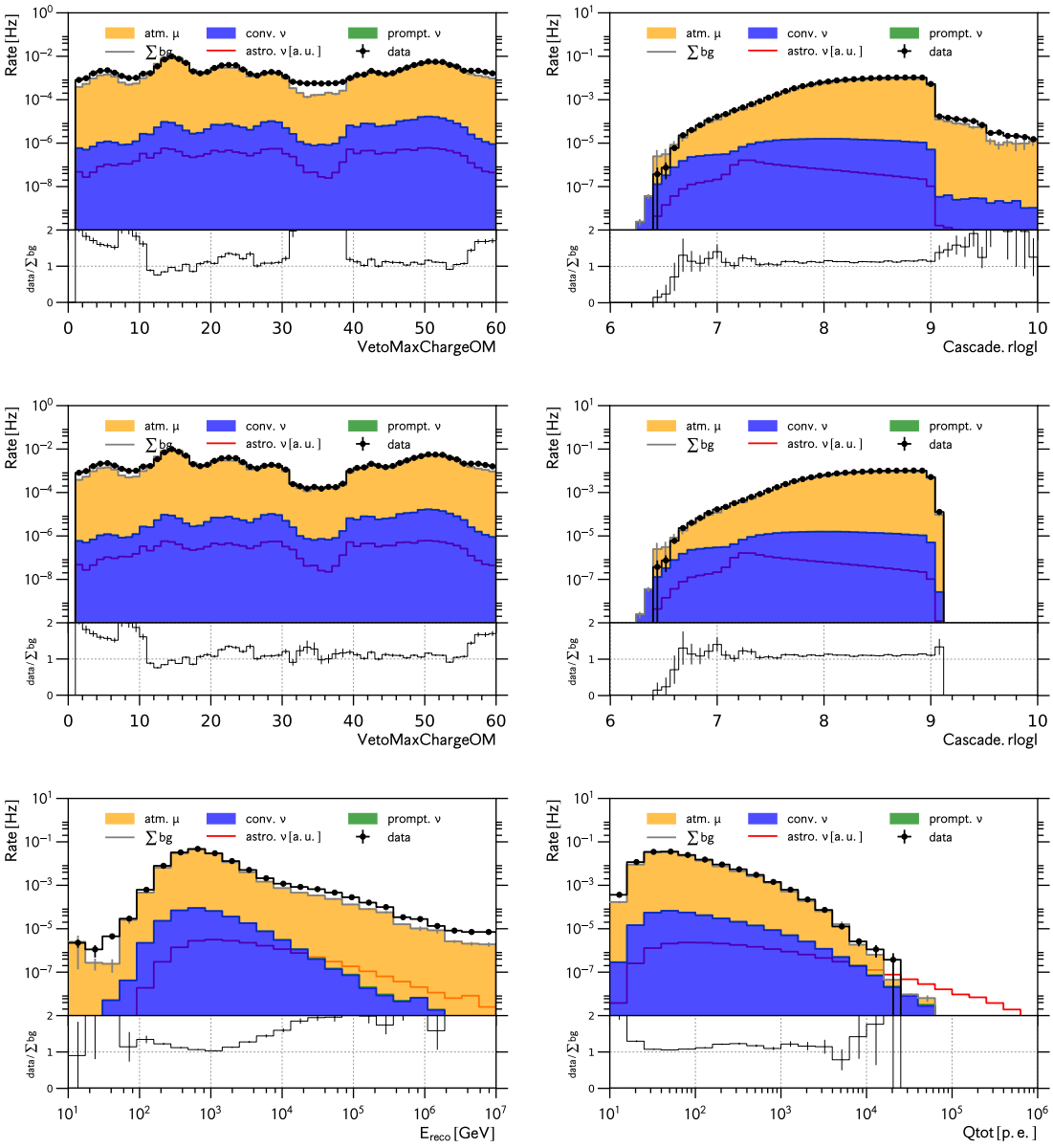


FIGURE 3.2: Relevant Observable Distributions - Top: VetoMaxChargeOM (left) and Cascade.rlogl (right) after Cascade-Level3 (SC); 2nd row: VetoMaxChargeOM (left) and Cascade.rlogl (right) after Level-4A selection; Bottom: Reconstructed energy and total PMT charge after Level-4A selection. data (10%) and MC (CORSIKA and NuGen) for 2011 configuration.

simulation prediction is observed. We studied this region of observable space in detail and found that most of the excess is identifiable by the quality parameter of the cascade reconstruction (cascade.rlogl from CascadeLlhVertexFit), that assigns large values to such events (smaller values correspond to higher quality, c.f. Sec. 3.1.1). This suggests a conditional selection. Similarly, we observed much improved data/mc agreement for dust-layer events with reconstructed (by Monopod) vertices within the dust-layer or larger energy deposits (as measured in total charge). We define the following criterion:

Rate in Year	**2015** [mHz]	**2014** [mHz]	**2013** [mHz]	**2012** [mHz]	2011 [mHz]
Data (10%)	$\mathbf{150.1 \pm 0.2}$	$\mathbf{150.8 \pm 0.2}$	$\mathbf{150.3 \pm 0.2}$	$\mathbf{188.3 \pm 0.3}$	142.6 ± 0.2
CORSIKA	n/a	n/a	n/a	$\mathbf{181.1 \pm 0.3}$	128.3 ± 0.5

TABLE 3.2: After Level-4A selection: Passing rates of experimental data for data taking periods 2011-2015 in comparison to estimated passing rates from Monte Carlo simulations (CORSIKA). n/a ≡ simulation not available.

Keep event if

$$\begin{aligned} &\text{VetoMaxChargeOM} \leq 30 \quad \text{or} \quad \text{VetoMaxChargeOM} \geq 39 \\ &\text{or} \quad \text{Cascade.rlogl} < 7.5 \\ &\text{or} \quad -200\,\text{m} < \text{Monopod.z} < -50\,\text{m} \\ &\text{or} \quad \log_{10}\left(Q_{tot}^{HLC}/p.e.\right) > 200.0 \end{aligned} \tag{3.2}$$

otherwise reject it.

While studying the cascade reconstruction quality parameter cascade.rlogl we observed a subdominant tail towards large values (bad quality), shown in Fig. 3.2 (top right). We have no interest in such events and remove them entirely.

Keep event if

$$\text{Cascade.rlogl} < 9.1 \tag{3.3}$$

otherwise reject it.

Both conditions (eqs. (3.2) **and** (3.3)) have to be satisfied simultaneously for an event to pass this stage of the pre-selection (logical and), which will subsequently be referred to as Level-4A filter criteria. Fig. 3.2 (bottom) shows the reconstructed energy (left) and total charge (right) distributions for all events that are retained by Level-4A. The spectra appear unchanged in comparison to Level-3 (Fig. 3.1), because of the small overall contribution from dust-layer events to the entire sample. However the effect of the cuts is clearly visible in the VetoMaxCharge variable (middle, left): The dust-layer excess is removed. The Level-4A passing rates are given in Tab. 3.2.

3.1.4 Level 4B: Containment

The cascade-Level 3 filter stream (contained branch) defines containment loosely based on charge criteria (Sec. 2.3). Event reconstruction and cascade identification performs better for events with contained vertices than for events with vertices beyond the detector boundaries. Hence such events suffer from larger background contamination, in particular muon bundles with large energy depositions outside but near the detector. Fig. 3.3 (top) shows the reconstructed vertex z-coordinate (left) as well as XYscale (right). The data exceeds the simulation prediction near the top ($z = 500\,\text{m}$) and bottom ($z = -500\,\text{m}$). Similarly a large increasing excess is observed for events that are reconstructed beyond the detector boundary in the XY-plane (XYscale > 1.0). We have no interest in such events.

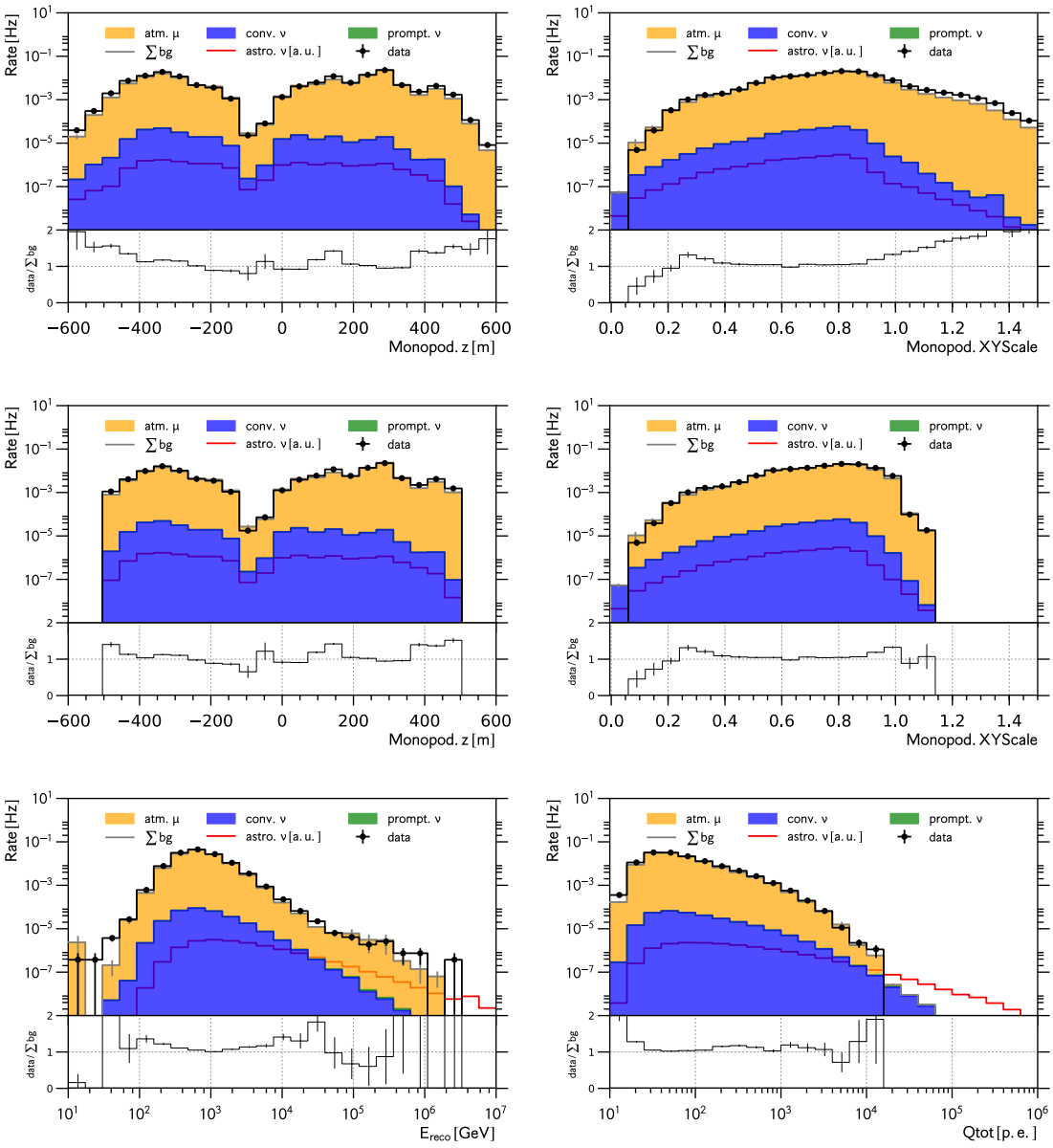


FIGURE 3.3: Relevant Observable Distributions - Top: Monopod.Z (left) and Monopod.XYScale (right) after Level-4A selection; 2nd row: Monopod.Z (left) and Monopod.XYScale (right) after Level-4B selection; Bottom: Reconstructed energy and total PMT charge after Level-4B selection. data (10%) and MC (CORSIKA and NuGen) for 2011 configuration.

There are dedicated analyses dealing with partially contained cascades, e.g. [154][156]. Hence we decided to remove the vast majority of such events from the sample. The filter criteria read:

Keep event if

$$\begin{aligned} &-500\,\mathrm{m} < \mathrm{Monopod.z} < 500\,\mathrm{m} \quad \text{and} \\ &\{\mathrm{Monopod.XYscale} < 1.0 \quad \text{or} \\ &(\mathrm{Monopod.XYscale} < 1.1 \quad \text{and} \quad \mathrm{Cscd.rlogl} < 7.6)\} \end{aligned} \tag{3.4}$$

otherwise reject it.

The last conditions retains a small contribution from events that appear with high quality cascade-like light deposition and are reconstructed not too far from the detector boundary. These events contain information about cascade-like background near the detector boundary and are well described by simulations, after the criteria from eq. (3.4) are applied. This is shown in Fig. 3.3 (middle, right). The reconstructed energy spectrum as well as the total charge distribution after applying the containment criteria are shown in Fig. 3.3 (bottom). The high energy ($E > 10\,\mathrm{TeV}$) tail of mis-reconstructed events, we observed earlier (c.f. Sec. 3.1.2), appears much reduced and the excess is less significant. The passing rates of the Level-4B pre-selection is given in Tab. 3.3.

Rate in Year	**2015** [mHz]	**2014** [mHz]	**2013** [mHz]	**2012** [mHz]	2011 [mHz]
Data (10%)	$\mathbf{134.3 \pm 0.2}$	$\mathbf{134.7 \pm 0.2}$	$\mathbf{134.2 \pm 0.2}$	$\mathbf{168.4 \pm 0.2}$	127.6 ± 0.2
CORSIKA	n/a	n/a	n/a	$\mathbf{165.9 \pm 0.3}$	118.7 ± 0.5

TABLE 3.3: After Level-4B selection: Passing rates of experimental data for data taking periods 2011-2015 in comparison to estimated passing rates from Monte Carlo simulations (CORSIKA). n/a ≡ simulation not available.

3.1.5 Level 4C: Anti-Top/Bottom

While Level-4B improved the data/mc agreement quite a lot, the data still do not appear to match the simulation prediction at very top (bottom) of the detector, as shown in Fig. 3.3 (2nd row, left). Since at low energies the muon background rate decreases as a function of depth, i.e. as a function of the thickness of the ice-shield, the excess at the top is believed to be due to low energy muon background. In principle such an excess could also be related to mis-modeling of the local optical properties of the ice. Unfortunately this cannot be checked, because we have no ability to re-simulated CORSIKA background events for alternative assumptions about the detector properties. Since we observe the same top/bottom data excess in the VetoMaxChargeDOM variable, a reconstruction effect appears unlikely, see Fig. 3.4 (top right). None of these possibilities are desirable, and thus we will tighten the containment criteria corresponding to a 140 m (50 m) rejection region at the top (bottom) of the detector. Here we use the reconstructed vertex position (z-coordinate) by Monopod. If Cascade.rlogl indicates poor reconstruction quality, then additional cuts based on the VetoMaxChargeOM are applied making this top/bottom rejection robust against reconstruction failures. The selection criteria are given in eq. (3.5).

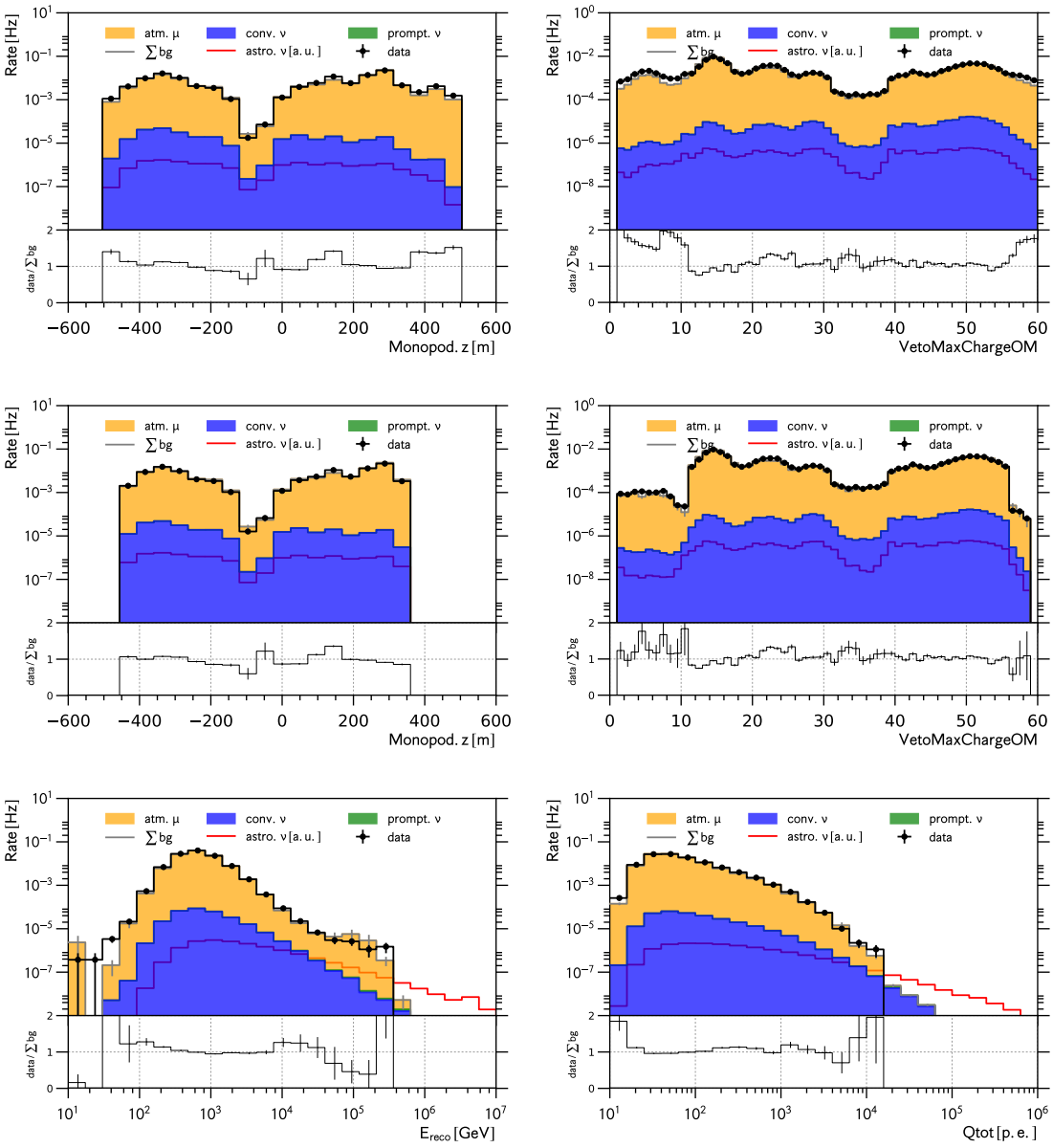


FIGURE 3.4: Relevant Observable Distributions - Top: Monopod.Z (left) and VetoMaxChargeOM (right) after Level-4B selection; 2nd row: Monopod.Z (left) and VetoMaxChargeOM (right) after Level-4C selection; Bottom: Reconstructed energy and total PMT charge after Level-4C selection. data (10%) and MC (CORSIKA and NuGen) for 2011 configuration.

Keep event if

$$\begin{aligned} &-450\,\text{m} < \text{Monopod.z} < 360\,\text{m} \\ &\text{and} \quad (10 < \text{VetoMaxChargeOM} < 56 \quad \text{or} \quad \text{Cascade.rlogl} < 7.5) \\ &\text{and} \quad \text{VetoMaxChargeOM} < 59 \text{ (removes 2 DOM-layers at bottom)} \end{aligned} \tag{3.5}$$

otherwise reject it.

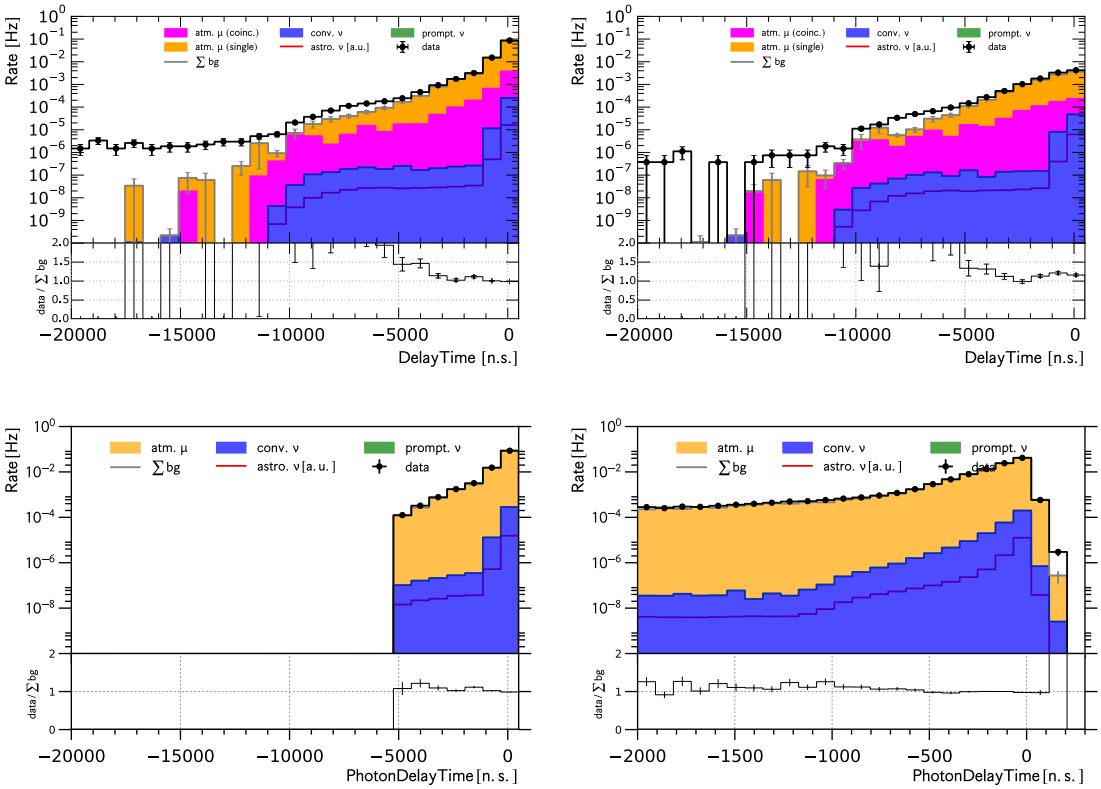


FIGURE 3.5: Relevant Observable Distributions - Top: DelayTime (left) and DelayTime ($+Q_{tot} > 200$ p.e.) (right) after Level-4C selection; Bottom: DelayTime (left) and DelayTime (zoom) (right) after Level-4D selection. data (10%) and MC (CORSIKA and NuGen) for 2011 configuration.

Fig. 3.4 (2nd row) shows the depth related variables after application of these Level-4C rejection criteria. The energy spectrum and total charge distribution are shown in the bottom row. The high energy excess, discussed before, essentially disappeared. The Level-4C passing rates are given in Tab. 3.4.

Rate in Year	**2015** [mHz]	**2014** [mHz]	**2013** [mHz]	**2012** [mHz]	2011 [mHz]
Data (10%)	$\mathbf{113.5 \pm 0.2}$	$\mathbf{114.2 \pm 0.2}$	$\mathbf{113.6 \pm 0.2}$	$\mathbf{142.4 \pm 0.3}$	108.6 ± 0.2
CORSIKA	n/a	n/a	n/a	$\mathbf{147.6 \pm 0.3}$	107.9 ± 0.5

TABLE 3.4: After Level-4C selection: Passing rates of experimental data for data taking periods 2011-2015 in comparison to estimated passing rates from Monte Carlo simulations (CORSIKA). n/a ≡ simulation not available.

3.1.6 Level 4D: Removing Large Delay Times

The last pre-selection filter removes events with very large (negative) photon delay times (Photon Delay Time, c.f. Sec. 3.1.1), which form a subdominant tail that is not well described by our simulation. This is shown in Fig. 3.5 (top) for all events (left) and events

Rate in Year	**2015** [mHz]	**2014** [mHz]	**2013** [mHz]	**2012** [mHz]	2011 [mHz]
Data (10%)	$\mathbf{113.5 \pm 0.2}$	$\mathbf{112.8 \pm 0.2}$	$\mathbf{112.3 \pm 0.2}$	$\mathbf{140.8 \pm 0.3}$	107.5 ± 0.2
CORSIKA	n/a	n/a	n/a	$\mathbf{146.4 \pm 0.3}$	107.4 ± 0.5

TABLE 3.5: After Level-4D selection: Passing rates of experimental data for data taking periods 2011-2015 in comparison to estimated passing rates from Monte Carlo simulations (CORSIKA). n/a ≡ simulation not available.

with more than 200 p.e. total charge (right). Large negative delay times are expected for coincident events, events with more than one atmospheric muon track from several uncorrelated air showers. Our simulation allows for the prediction of the coincident background component (pink histogram). While the relative contribution from coincident background increases towards larger negative delay times, the absolute normalization is too small to explain the data excess. We investigated some of these data events manually with the event viewer and found them to be consistent with background. Thus we decided to remove the tail area from the analysis. The criterion reads:

Keep event if

$$\begin{aligned} &-3000\,\text{ns} < \text{Photon Delay Time} < 125\,\text{ns} \quad \text{or} \\ &(-5000\,\text{ns} < \text{Photon Delay Time} \leq 3000\,\text{ns} \\ &\quad \text{and} \quad 0.35 < \text{Monopod.XYscale} < 0.8) \end{aligned} \tag{3.6}$$

otherwise reject event.

We can extent the acceptance region down to −5000 ns by tightening the containment cut (XY-plane) to maintain data/mc agreement. The distribution of the delay times after these Level-4D selections is shown in Fig. 3.5 (bottom) and the corresponding passing rates can be found in Tab. 3.5.

Since this concludes our preselection, it is worth revisiting the impact of the selection criteria on the reconstructed energy spectrum. Fig. 3.6 compares the reconstructed energy spectrum without any pre-selection (top) to the one obtained after applying the filter criteria defined in Level-4 A-D (bottom). The test-data from the year 2011, used to develop the criteria, is shown in the left column. The criteria generalize well, as demonstrated by the good data/mc agreement observed for the test-data from the year 2012 (right column). While we did not strongly reduce the overall background rate, the pre-filter criteria remove the majority of the mis-reconstructed high-energy background events thereby improving the data/mc agreement to the observed level. Since we suppressed miss-reconstructed events, no strong impact on the total charge distribution is expected. The corresponding distributions after Level4 A-D are shown in Fig. 3.7 (bottom) and can be compared to the ones at the starting point of the selection, Fig. 3.1 (bottom). The shape of the distributions remains largely unchanged.

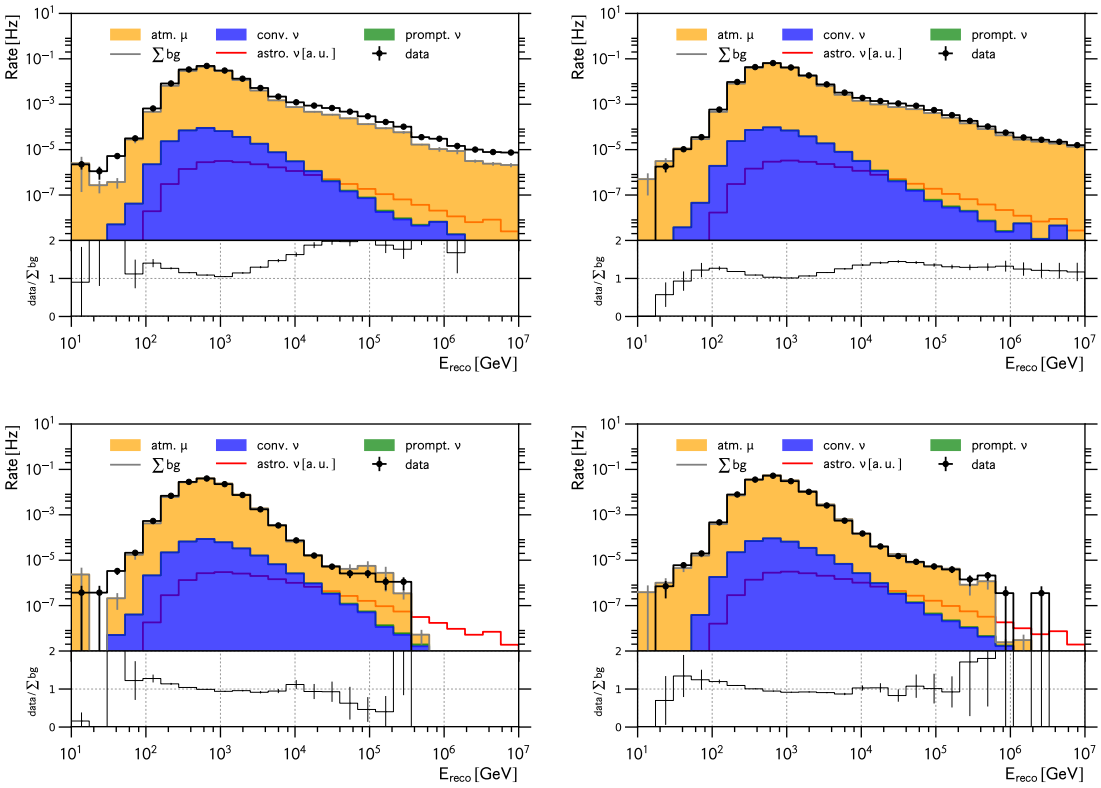


FIGURE 3.6: Before/After comparison of reconstructed energy distributions - Top: Reconstructed energy for 2011 (left) and 2012 (right) configuration before any L4 selections; Bottom: Reconstructed energy for 2011 (left) and 2012 (right) configuration after Level-4D selection.

Rate in Year	**2015** [μHz]	**2014** [μHz]	**2013** [μHz]	**2012** [μHz]
ν_e (conv.)	c.f. 2013	c.f. 2013	$\mathbf{54.51 \pm 0.06}$	$\mathbf{55.50 \pm 0.06}$
ν_μ (conv.)	c.f. 2013	c.f. 2013	$\mathbf{253.5 \pm 0.7}$	$\mathbf{259.5 \pm 0.8}$
ν_e (prompt)	c.f. 2013	c.f. 2013	$\mathbf{0.9176 \pm 0.0006}$	$\mathbf{0.9283 \pm 0.0006}$
ν_μ (prompt)	c.f. 2013	c.f. 2013	$\mathbf{0.2776 \pm 0.0007}$	$\mathbf{0.2827 \pm 0.0007}$
ν_e (astro)	c.f. 2013	c.f. 2013	$\mathbf{5.707 \pm 0.004}$	$\mathbf{5.770 \pm 0.004}$
ν_μ (astro)	c.f. 2013	c.f. 2013	$\mathbf{2.041 \pm 0.004}$	$\mathbf{2.087 \pm 0.004}$
ν_τ (astro)	c.f. 2013	c.f. 2013	$\mathbf{4.229 \pm 0.006}$	$\mathbf{4.275 \pm 0.006}$

TABLE 3.6: After Level-4D selection: Estimated neutrino passing rates for data taking periods 2012-2015 based on NuGen simulation for the 2012 and 2013 configurations.

3.1.7 Level 4: Summary

The Level 4 selection criteria reduce the total data volume (muon background rate) by $\sim 27\%$, see Tab. 3.5 (L4D) and Tab. 3.1 (L3). The estimated contribution from neutrinos is $\sim 0.2\%$ and is given in Tab. 3.6 for the different neutrino flavors and production mechanisms. Thus far, we have not discussed the influence of the selection criteria on the neutrino component, because the criteria are mostly either geometric in nature or require high reconstruction quality, and are not meant to distinguish further between muon

background and neutrino signal. Fig. 3.7 (top) shows the neutrino selection efficiency of the Level 4 criteria, compared to the starting point of this analysis, as a function of neutrino energy. Independent of neutrino flavor the selection efficiency is $\sim 95\%$ for low energy neutrinos ($E_\nu = 1\,\mathrm{TeV}$). At higher energies ($E_\nu = 1\,\mathrm{PeV}$) it is slightly lower ($\sim 85\%$) because of the containment criteria (L4-B and L4-C). Since some of the criteria are more constraining for events that appear less cascade-like (low cascade fit quality), the high energy efficiency for muon neutrinos is slightly lower than that of electron and tau neutrinos.
Finally we show that the selection criteria developed using test-data and CORSIKA simulation for the 2011 detector configuration (left column) generalize well to the 2012 configuration (right column) for all relevant observables, see Figs. 3.7 (bottom), 3.8, 3.9 and 3.10. All of these variables will be used to build the gradient boosted decision tree model, introduced in the next chapter.

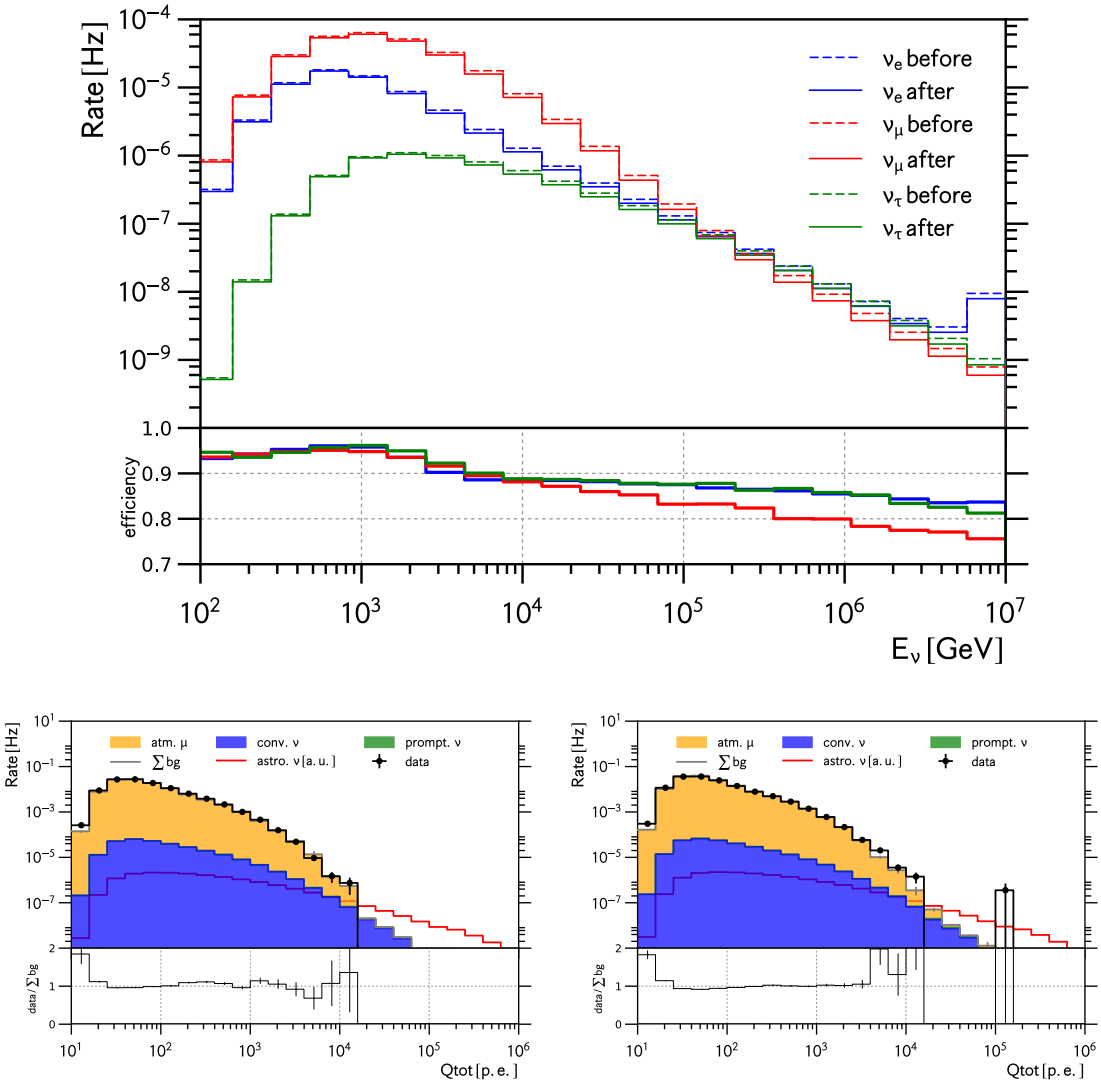


FIGURE 3.7: Top: Neutrino selection efficiencies as function of neutrino energy of all Level-4 (A-D) criteria with respect to Cascade-Level3 (SC) filter. Bottom: Distribution of total PMT charge for 2011 (left) and 2012 (right) configurations.

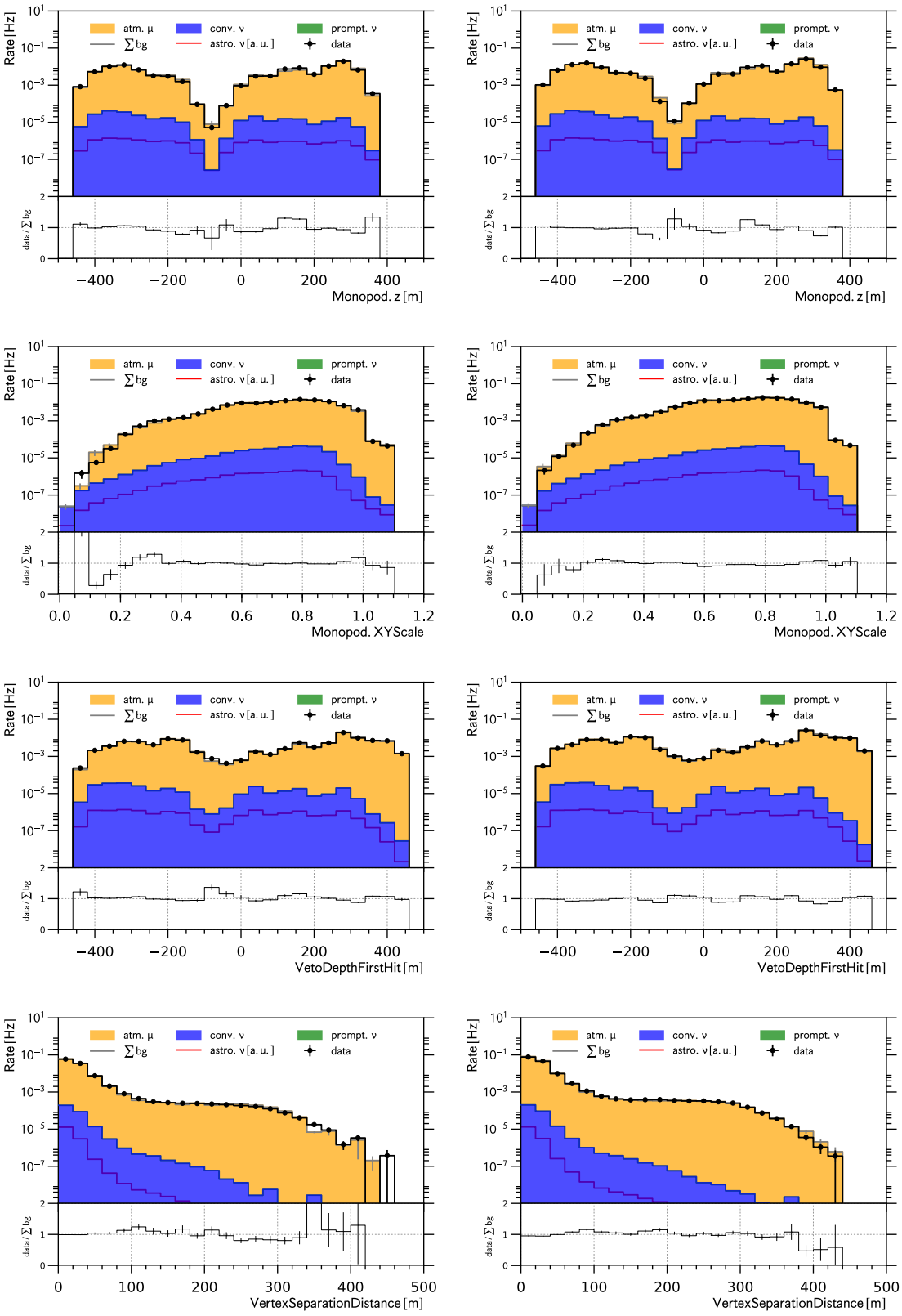


FIGURE 3.8: After Level-4D selection: Observable distributions for data periods 2011 (left) and 2012 (right).

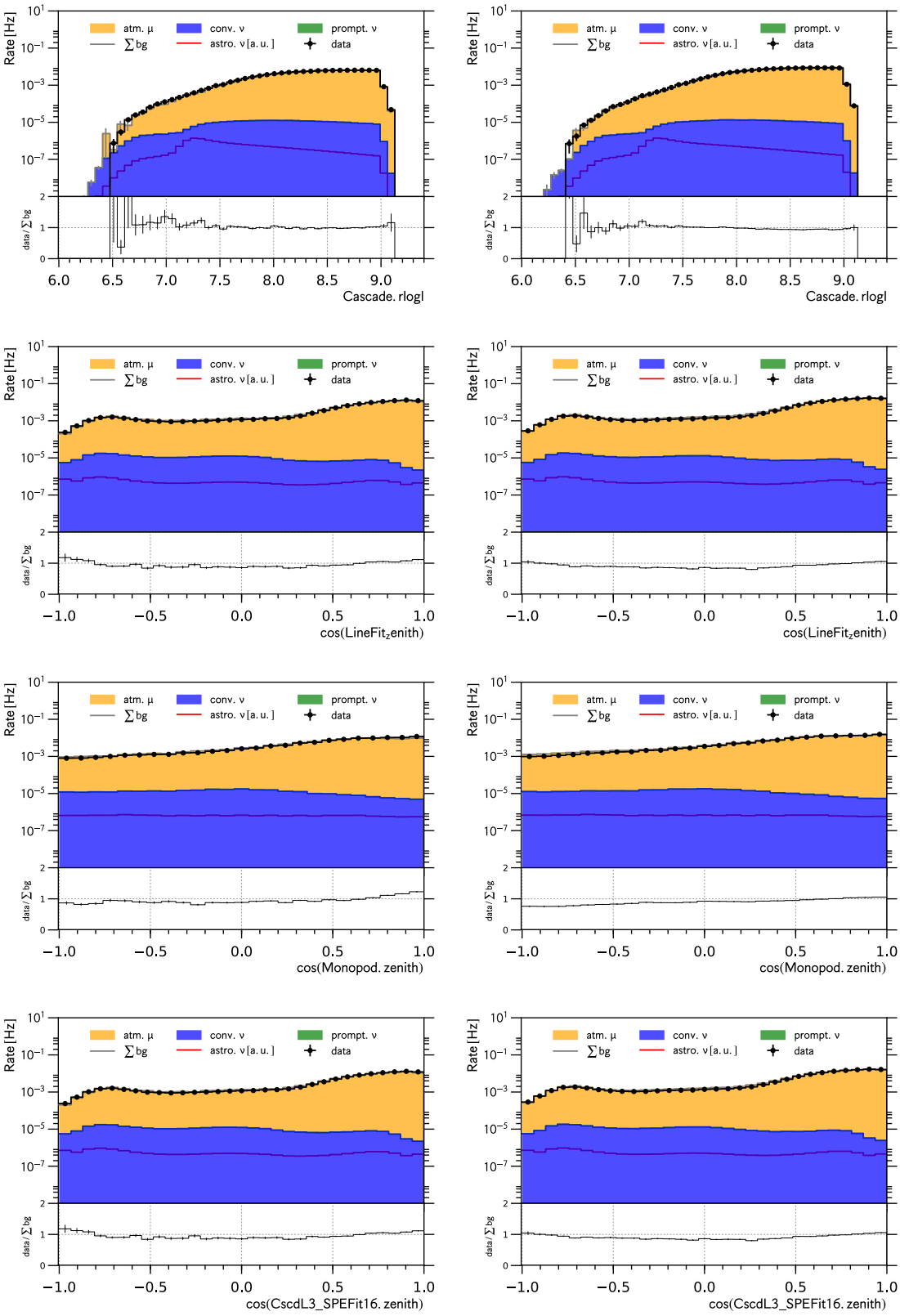


FIGURE 3.9: After Level-4D selection: Observable distributions for data periods 2011 (left) and 2012 (right).

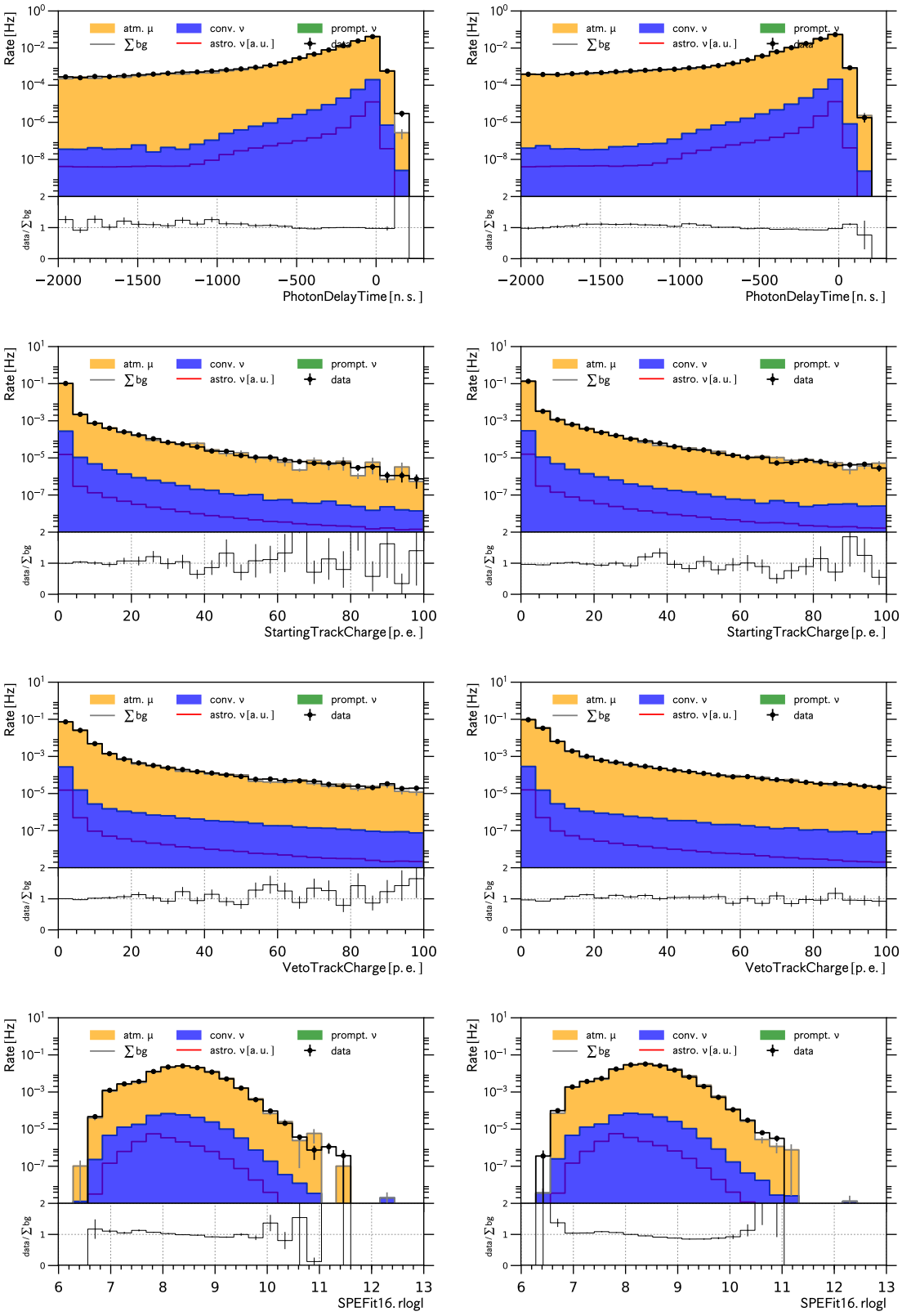


FIGURE 3.10: After Level-4D selection: Observable distributions for data periods 2011 (left) and 2012 (right).

3.2 Event Identification and Background Rejection: Machine Learning

In order to achieve our goal of isolating an at least 90% pure sample of neutrinos we need to suppress the atmospheric muon background by several orders of magnitude, while keeping the neutrino selection efficiency as large as possible. At the same time we are interested in distinguishing between neutrino flavors, especially in the identification and separation of starting tracks (muon neutrino charged-current interactions) from the cascade-like signal events. We can do all of the above in one step by means of multi-class classification methods, rooted in machine learning. In particular we will show that **Gradient Boosted Decision Trees** (GBDT) [157] perform well in the context of our problem. We will give a brief introduction into the method in Sec. 3.2.1. Following [158][159], we define a statistical (non-parametric) model and will fit it to the data (MC simulations of the actual physical processes) by means of maximum likelihood optimization. The method of maximum likelihood is introduced in Sec. 5.3 in the context of fitting astrophysical neutrino flux models to the final sample. The concept is identical and thus we refer to Sec. 5.3 for general definitions.
In this section we demonstrate the performance of the GBDT model and define the final selection criteria based on its predictions. The resulting selection will be combined with the dedicated high-energy neutrino selection from [153], as discussed in Sec. 3.3.

3.2.1 Introduction: Gradient Boosted Decision Trees

Formulating the Statistical Model

Suppose we collect data a dataset $\mathbf{D} = \{(Y_i, \mathbf{X}_i)\,;\, 1 \leq i \leq N\}$ with N events. For each event i we would like to know the value Y_i, but we only measured the values of the observables $\mathbf{X_i}$. Thus we have to predict the latent variable Y_i from the observations $\mathbf{X_i}$. This requires to specify a statistical model $f_{Y|\mathbf{X}}\left(Y \,|\, \mathbf{X},\, \boldsymbol{\theta}\right)$ of the data generating process relating Y_i and $\mathbf{X_i}$.

$$Y|\boldsymbol{X} \sim f_{Y|\mathbf{X}}\left(Y \,|\, \mathbf{X},\, \boldsymbol{\theta}\right) \tag{3.7}$$

where $\boldsymbol{\theta}$ are unknown model parameters. For now consider Y_i to be a continuous random variable. We could assume a normal distribution with variance σ^2 to define a *regression problem* for its mean value $\mu = \mu\left(\mathbf{x},\, \boldsymbol{\theta}\right) = f\left(\mathbf{x},\, \boldsymbol{\theta}\right)$ as a function of the observed value $\mathbf{x}$ of $\mathbf{X}$ with parameters $\boldsymbol{\theta}$:

$$Y|\boldsymbol{X} \sim \mathrm{N}\left(Y \,|\, f\left(\boldsymbol{X},\, \boldsymbol{\theta}\right),\, \sigma^2\right) \tag{3.8}$$

If we know the function $f\left(\mathbf{x},\, \boldsymbol{\theta}\right)$ and true parameter values $\boldsymbol{\theta_0}$, prediction of Y_i would be straightforward, e.g. replacing the unobserved value Y_i with the expectation Y_i^e from the model: $Y_i^e = E\left(Y_i \,|\, \mathbf{X_i}\right) = f\left(\boldsymbol{X_i},\, \boldsymbol{\theta_0}\right)$[4].
If we do not know the true parameter values $\boldsymbol{\theta_0}$, but we know the functional form

[4]We use $E(\cdot)$ to denote the average: $E(X) = \int x \cdot f_X\left(x\right)\,\mathrm{dx}$.

of $f(\boldsymbol{x}, \boldsymbol{\theta})$, then we could estimate $\boldsymbol{\theta_0}$ from an independent *training-dataset* $\mathbf{D_{train}} = \{(Y_i', \mathbf{X}_i')\,;\, 1 \leq i \leq N'\}$. One would fit the model, eq. (3.8), to the training-set for example using the method of maximum likelihood (c.f. Sec. 5.3) and predict Y_i as follows: $Y_i^p = f(\boldsymbol{X_i}, \hat{\boldsymbol{\theta}})$. Here we used $\hat{\boldsymbol{\theta}}$ to denote the corresponding best-fit value. For the regression model (3.8), this leads to the so-called chi-squared solution:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\mathbf{argmin}} \sum_{i=1}^{N'} \left\{Y_i' - f\left(\mathbf{X_i'}, \boldsymbol{\theta}\right)\right\}^2 \tag{3.9}$$

The sum is often referred to as *squared-error-loss*, when normalized by the total number of samples.

Now consider the *binary classification problem*, i.e. we are interested in predicting *discrete* variables $Y_i = \{0, 1\}$, where we define $Y_i = 1$ ($Y_i = 0$) as signal (background) event. The statistical model would assume a Bernoulli probability mass function (Binomial probability mass function with $n = 1$) with success probability $p = p(\mathbf{x}, \boldsymbol{\theta}) \equiv f(\mathbf{x}, \boldsymbol{\theta})$ depending on the observables $\mathbf{X}$ and parameters $\boldsymbol{\theta}$. The model reads

$$Y|\boldsymbol{X} \sim \text{binom}\left(Y \,|\, p = f\left(\boldsymbol{X}, \boldsymbol{\theta}\right), n = 1\right) \tag{3.10}$$

The (negative) log-likelihood function (c.f. Sec. 5.3) becomes

$$-\log L = -\sum_{i=1}^{N'} \left\{Y_i' \log f\left(\boldsymbol{X_i'}, \boldsymbol{\theta}\right) + (1 - Y_i') \log(1 - f\left(\boldsymbol{X_i'}, \boldsymbol{\theta}\right))\right\} \tag{3.11}$$

and is often referred to as *log-loss* or *cross-entropy*, when normalized by the total number of samples. It is numerically inconvenient to work with $p(\mathbf{x}, \boldsymbol{\theta}) \equiv f(\mathbf{x}, \boldsymbol{\theta})$ since, being a probability, it is bounded between $0 \leq f(\mathbf{x}, \boldsymbol{\theta}) \leq 1$. This can be resolved by means of a *logistic transformation*

$$\begin{aligned} g\left(\boldsymbol{x}, \boldsymbol{\theta}\right) = \text{logit}\left(p\left(\mathbf{x}, \boldsymbol{\theta}\right)\right) &\equiv \log\left(\frac{f\left(\mathbf{x}, \boldsymbol{\theta}\right)}{1 - f\left(\mathbf{x}, \boldsymbol{\theta}\right)}\right) \\ f\left(\mathbf{x}, \boldsymbol{\theta}\right) &= \frac{1}{1 + \exp\left(-g\left(\mathbf{x}, \boldsymbol{\theta}\right)\right)} \end{aligned} \tag{3.12}$$

Inserting eq. (3.12) in eq. (3.11) yields the following best-fit solution for the parameter vector $\boldsymbol{\theta}$

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\mathbf{argmin}} \sum_{i=1}^{N'} \log\left\{1 + \exp\left(g\left(\boldsymbol{x}, \boldsymbol{\theta}\right)\right)\right\} - Y_i' g\left(\boldsymbol{x}, \boldsymbol{\theta}\right) \tag{3.13}$$

We should point out that this requires a parametric form for the logarithm of the binomial odds ratio $g(\boldsymbol{x}, \boldsymbol{\theta})$. The prediction for Y_i becomes:

$$Y_i^p = E\left(Y_i|\mathbf{X}\right) = p\left(\mathbf{X}, \boldsymbol{\theta}\right) \tag{3.14}$$

where we need to replace the unknown parameters $\boldsymbol{\theta}$ by the best-fit values $\hat{\boldsymbol{\theta}}$.

$$\hat{p}\left(\mathbf{x_i}\right) = p\left(\mathbf{x_i},\, \hat{\boldsymbol{\theta}}\right) = \frac{1}{1+\exp\left(-g\left(\boldsymbol{x_i},\, \hat{\boldsymbol{\theta}}\right)\right)} \tag{3.15}$$

Thus the prediction for the unobserved quantity is the estimated success (signal event) probability of the binomial model from its best-fit solution.

In this work we are interested in the *multi-class classification problem* with $N_c = 3$ classes: atmospheric muon background, neutrino-induced starting tracks and neutrino-induced cascades. The model follows from the previous one by replacing the binomial model with a multinomial one:

$$\begin{aligned} &\mathbf{Y} \,|\, \mathbf{X} \sim \text{multinom}\left(\mathbf{Y} \,|\, \mathbf{p} = \mathbf{f}\left(\mathbf{X}, \boldsymbol{\theta}\right),\, n = 1\right) \\ &\mathbf{Y} = \left(Y_1,\, ...,\, Y_{N_c}\right), \quad Y_i \in \{0, 1\}, \quad \sum Y_i = n = 1 \\ &\mathbf{p} = \left(p_1,\, ...,\, p_{N_c}\right), \quad \sum p_i = 1 \end{aligned} \tag{3.16}$$

The (negative) log-likelihood function (c.f. Sec. 5.3) becomes

$$\begin{aligned} -\log L = -\sum_{i=1}^{N'} \sum_{k=1}^{N_c} I\left(\left(\mathbf{Y}'_i\right)_k = 1\right) \times \log f_k\left(\mathbf{X}'_\mathbf{i}, \boldsymbol{\theta}\right) \\ I\left(\left(\mathbf{Y}'_i\right)_k = 1\right) = \begin{cases} 1 & \text{if } \left(\mathbf{Y}'_i\right)_k = 1 \\ 0 & \text{if } \left(\mathbf{Y}'_i\right)_k = 0 \end{cases} \end{aligned} \tag{3.17}$$

and is often referred to as *multi-class log-loss*, when normalized by the total number of samples.
For reasons of numerical convenience it will again be helpful to transform $\mathbf{f}\left(\mathbf{X}, \boldsymbol{\theta}\right)$ (vector-valued) onto an unbounded scale. For the binomial problem, a suitable transformation was the logit-transformation eq. (3.12). Its generalization to the multinomial problem (vector of probabilities), the *softmax transformation* [159] reads

$$p_k\left(\mathbf{x},\, \boldsymbol{\theta}\right) = \frac{\exp\left(g_k\left(\mathbf{x},\, \boldsymbol{\theta}\right)\right)}{\sum_{i=1}^{N_c} \exp\left(g_i\left(\mathbf{x},\, \boldsymbol{\theta}\right)\right)} \tag{3.18}$$

Inserting eq. (3.18) in eq. (3.17) yields the following best-fit solution for the parameter vector $\boldsymbol{\theta}$ [159]

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\mathbf{argmin}} \sum_{i=1}^{N'} - \left\{ \left(\sum_{k=1}^{N_c} I\left(\left(\mathbf{Y}'_i\right)_k = 1\right) \times g_k\left(\mathbf{X}'_\mathbf{i}, \boldsymbol{\theta}\right) \right) + \log\left(\sum_{k=1}^{N_c} \exp\left(g_k\left(\mathbf{X}'_\mathbf{i}, \boldsymbol{\theta}\right)\right) \right) \right\} \tag{3.19}$$

After having fit a parametric function $\mathbf{g}\left(\mathbf{x}, \boldsymbol{\theta}\right)$, we can, for each observation $\left(Y_i,\, \mathbf{X_i}\right)$ and each class, estimate the corresponding membership probability, i.e. N_c probabilities per

observation.

$$\hat{p}_k\left(\mathbf{x_i}\right) = p_k\left(\mathbf{x_i}, \hat{\boldsymbol{\theta}}\right) = \frac{\exp\left(g_k\left(\mathbf{x_i}, \hat{\boldsymbol{\theta}}\right)\right)}{\sum_{i=1}^{N_c}\exp\left(g_i\left(\mathbf{x_i}, \hat{\boldsymbol{\theta}}\right)\right)} \tag{3.20}$$

Thus far the discussion has been general. We have not made assumptions about the character of the function $\mathbf{g}\left(\mathbf{x}, \boldsymbol{\theta}\right)$. However, we required knowledge about its parametric form, mostly because this allowed to illuminate the problem from the perspective of parametric statistics and, consequently standard fitting procedures. In most problems (including ours) no closed form parametric expression for the functions $\mathbf{g}\left(\mathbf{x}, \boldsymbol{\theta}\right)$ exists. The discussion above still applies, but the problem becomes more involved. We now have to estimate/approximate the function $\mathbf{g}\left(\mathbf{x}, \boldsymbol{\theta}\right)$ itself. Methods from *non-parametric* statistics need to be added. Note, as we have demonstrated above, this task is a common one - no matter whether regression, binary classification or multi-class classification is concerned. Hence we will not distinguish between those tasks in the following sections.

Decision Trees

Decision Trees are a conceptually simple yet powerful method of approximating high dimensional functions. In one dimension they reduce to piecewise constant functions. Decision Trees, $T\left(\mathbf{x}, \mathbf{c}, \mathbf{R}\right)$, are derived from partitioning the support (observable/feature/variable space) of the function via recursive binary splitting into rectangular regions R_i. The approximation $\hat{g}\left(\mathbf{x}\right)$ of the function $g\left(\mathbf{x}\right)$ reads

$$\begin{aligned} \hat{g}\left(\mathbf{x}\right) = T\left(\mathbf{x}, \hat{\mathbf{c}}, \hat{\mathbf{R}}\right) &= \sum_{i=1}^{M} \hat{c}_i \times I\left(\mathbf{x} \in \hat{R}_i\right) \\ I\left(\mathbf{x} \in R_i\right) &= \begin{cases} 1 & \text{if } \mathbf{x} \in R_i \\ 0 & \text{if } \mathbf{x} \notin R_i \end{cases} \end{aligned} \tag{3.21}$$

The complexity and thus the flexibility of this tree model obviously increases with the total number of regions $M = 2^K$. Here we define K as the *depth of the tree*, i.e. the total number of recursive splits. We will not discuss different implementations and use cases of decision tree models here and refer to standard references [159][160] instead. It turns out that a model built from a superposition of many rather shallow (few splits) decision trees is vastly superior to a single, deep (many splits) decision tree model in terms of performance, i.e. its predictive accuracy, for many problems [159]. This motivated a procedure called *Boosting*[5].

[5]Boosting is in principle not tied to decision trees. Here we only discuss it in the context of decision trees.

Gradient Boosting

Boosting refers to the idea of approximating a function $g(\mathbf{x})$ by its expansion $\hat{g}(\mathbf{x})$ into a set of suitable basis functions $B(\mathbf{x}, \gamma_i)$.

$$\hat{g}(\mathbf{x}) = \sum_{i=1}^{M} \hat{\alpha}_i B(\mathbf{x}, \hat{\boldsymbol{\gamma}}_i) \tag{3.22}$$

where the values of the coefficients α_i and $\boldsymbol{\gamma}_i$ are assumed to have been determined from a suitable training-dataset. In our case we will use decision trees as a basis - thus the term Boosted Decision Trees.

$$\hat{g}(\mathbf{x}) = \sum_{i=1}^{M} \hat{\alpha}_i T\left(\mathbf{x}, \hat{\mathbf{c}}_\mathbf{i}, \hat{\mathbf{R}}_\mathbf{i}\right) \tag{3.23}$$

In principle we now have all necessary ingredients to fit the model. One would insert (3.23) into (3.19), perform the necessary minimization of the log-likelihood to find the best-fit values $\hat{\boldsymbol{\theta}} = \{\hat{\boldsymbol{\alpha}}, \hat{\mathbf{c}}, \hat{\mathbf{R}}\}$, that uniquely define the approximation $\hat{g}(\mathbf{x})$ of $g(\mathbf{x})$. Inserting into (3.20) would yield the desired probability estimates (predictions) $\hat{p}_k(\mathbf{x_i})$. In practice direct numerical minimization of eq. (3.19) for model (3.23) is impossible and a local optimization strategy, called *forward stagewise additive modeling* [159] is employed. The minimization is done by iteratively adding decision trees, one-by-one, that maximally reduce the log-likelihood compared to the result from the previous iteration. At iteration $m-1$ one solves

$$(\alpha_m, \mathbf{c}_m, \mathbf{R}_m) = \underset{\alpha, \mathbf{c}, \mathbf{R}}{\operatorname{argmin}} \{-\log L(\mathbf{Y}', g_{m-1}(\mathbf{x}) + \alpha T(\mathbf{x}, \mathbf{c}, \mathbf{R}))\} \tag{3.24}$$

$$g_m(\mathbf{x}) = g_{m-1}(\mathbf{x}) + \alpha_m T(\mathbf{x}, \mathbf{c}_m, \mathbf{R}_m) \tag{3.25}$$

either via *gradient descent* or *Newton's method*. The latter is implemented in the *XGBoost* library [152], which we chose for this work, because of its speed and accuracy[6].

To demonstrate the flexibility of Gradient Boosted Decision Trees as non-parametric function estimates, we fitted a simple regression model, eq. (3.8), to noisy observations from a true function (taken from [162]), which is constant over some part of its support, slowly varying and rapidly varying in other parts. The true function is shown in Fig. 3.11 (black) and compared to the best-fit GBDT approximation (yellow), which captures the behavior of the true function well. Also shown is an early estimate (red) from the boosting procedure (2% of total boosting iterations). This example illustrates one big advantage of GBDTs over other non-parametric methods: If the problem requires complexity in some parts of its observable space, the model does not tend to find such complexity in other regions where it might not be warranted. In particular, no "ringing" is observed in the central range where the true model is indeed constant.

[6] XGBoost has been awarded the "HEP meets ML" award for high performance at the Higgs boson machine learning challenge, hosted by the ATLAS experiment [161]. It also received the John M. Chambers Statistical Software Award of the American Statistical Association (2016).

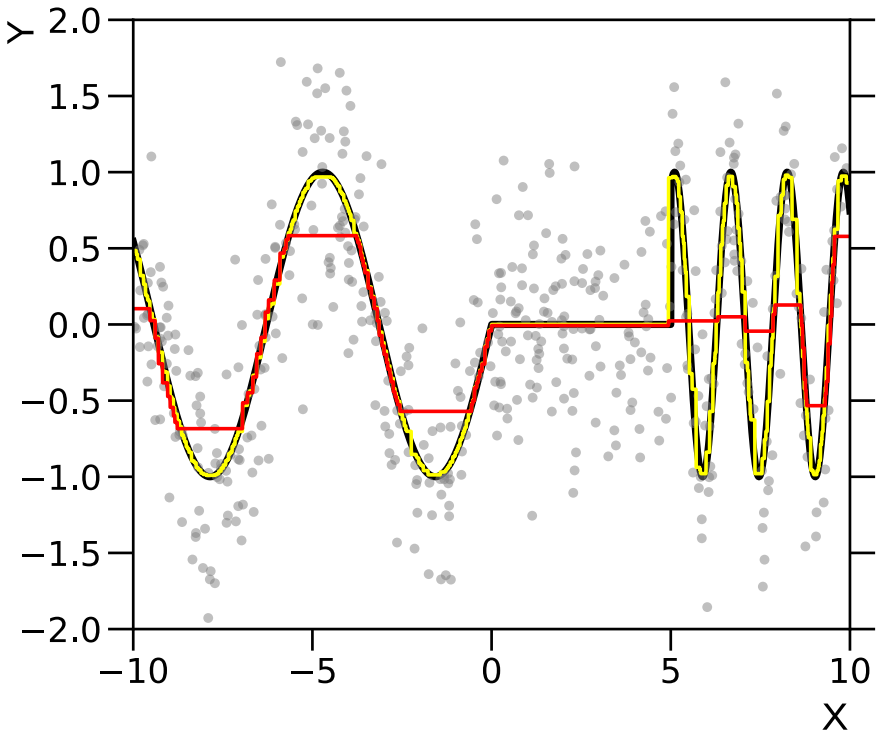


FIGURE 3.11: Example fit result for a non-parametric regression problem obtained with gradient boosted decision trees using XGBoost [152] (Figure and problem adapted from [162]).

Parameters that influence the behavior and performance of the fits and are not determined during the stagewise-optimization, eq. (3.24), are called *hyper-parameters*. The Boosting procedure introduces two such parameters: the *total number of boosting iterations* M (number of trees used to define the basis) and the *shrinkage*, introduced by [158]. The former is obvious but the latter needs explanation.
The (locally) optimal update of the model $g_{m-1}(\mathbf{x})$ at iteration $m-1$ is given by (3.25). [158] found that the stability of the final model always improves (in some cases significantly) if the stepsize is reduced by some *shrinkage factor* $\eta \ll 1$.

$$g_m(\mathbf{x}) = g_{m-1}(\mathbf{x}) + \eta \times \alpha_m T(\mathbf{x}, \mathbf{c}_m, \mathbf{R}_m) \tag{3.26}$$

However, this requires to increase the total number of boosting iterations M, since the log-likelihood reduction per step is reduced.
Another parameter corresponds to the individual decision trees. The *maximum depth of the tree* $T \geq 2$ controls the allowed complexity of each individual basis function. Typical values are $4 \leq T \leq 8$ [159]. Furthermore one might require a *minimum number of observations* (or total weight, in case of weighted observations) in each region R_m, thereby controlling their volume.
In addition XGBoost introduces *complexity penalization* or *regularization* for the individual trees into the log-likelihood function, eq. (3.17), during the stagewise fit. The

penalized log-likelihood function $-\log L_p$ reads:

$$-\log L_p = -\log L + \gamma T^* + \frac{1}{2}\lambda ||\mathbf{c}_m||_2^2 \tag{3.27}$$

where T^* is the number of terminal nodes of the tree ($T^* = 2^T$ if all branches of the tree reach maximum depth T) and $||\mathbf{c}_m||_2^2$ is the norm of the vector of constants determined for each tree region R_m. Large values of γ lead to a preference for small trees, while large values for λ penalize regions R_m with large values of c_m.

We have thus far omitted mentioning sample weights to maintain clean notation. As discussed in Sec. 2.5, IceCube simulations rely on importance sampling for computational efficiency. Thus all simulated events come with associated sample weights w_i. When fitting the GBDT model to simulated events, these weights are taken into account as follows:

$$-\log L = \sum_{i=1}^{N'} \log L_i(Y_i,\ f\left(\mathbf{X}_i, \boldsymbol{\theta}\right)) \longrightarrow \sum_{i=1}^{N'} w_i \log L_i(Y_i,\ f\left(\mathbf{X}_i, \boldsymbol{\theta}\right)) \tag{3.28}$$

3.2.2 Building the Model

In order to fit the GBTD multi-class model, eqs. (3.23) and (3.16), a suitable training dataset is necessary. During development of this event selection the CORSIKA simulation dataset corresponding to the 2011 detector configuration provided the largest number of simulated atmospheric muon background events for a recent detector configuration, and thus is used to model atmospheric muon background. After application of all pre-selection (Level 4 A-D, c.f. Sec. 3.1.7) criteria $\sim 300,000$ simulated muon background events are retained. To maintain approximate balance between the different classes, we add neutrino simulation for starting tracks (ν_μ with CC interactions) and cascades (ν_e with all interactions), each with similar total number of events, passing Level 4.
We use 13 observables with good agreement between data and simulation (after Level 4) to discriminate the three different classes of events. Hence $\dim \mathbf{X} = 13$. These variables are:

- $\log_{10}(Q_{tot}/p.e.)$
- Monopod.Z, XYscale and Zenith
- Depth of First Hit
- Cascade.rlogl and Track.rlogl
- LineFit.Zenith and SPEFit.Zenith
- Photon Delay Time
- Veto Track and Starting Track Charges
- Reconstructed Vertex Separation

Each variable is explained in Sec. 3.1.1 and shown in Figs. 3.8 to 3.10.

Thus far we have omitted discussing the *overfitting problem*. The method of maximum likelihood (c.f. Sec. 5.3) itself, while in a certain sense optimal for point estimation, does not provide a good criterion to distinguish between different models, as it essentially always prefers more complex models over simple ones. However the latter might generalize better to new and independent datasets. Here we define "overfitting" loosely as non-optimal model choice in favor of an overly complex model, thereby impeding its performance on independent data after the fit. We would like to avoid that and aim for maximal performance on independent data.
In order to select a model, which in our context means finding a suitable set of values for the GBDT hyper-parameters, introduced above, we need to define a metric. However no metric exists that could blindly be optimized. For example high energy cascade events are more "valuable" than low energy ones, since they contain a higher fraction of astrophysical neutrinos but low energy neutrinos are more numerous and thus provide betters statistics. Some muon background is acceptable at low energies, but interpreted as detrimental at high energies. Predictions for data and simulation should agree well. The model should perform similarly for the different detector seasons and so forth. Therefore we decided to first perform the optimization based on a simple, standard criterion, the multi-class log-loss (eq. (3.17)) (normalized, negative multinomial log-likelihood, "mlogloss"), before verifying manually using MC simulations and the 10% exp. test dataset that the corresponding model satisfies the desiderata from above.
Since we are interested in measuring predictive performance, we evaluate the multiclass-logloss criterion, eq. (3.17), on independent *validation data* (CORSIKA and neutrino simulation) that is not used during the fitting process, i.e. disjoint from the training dataset. Since, as discussed, the available CORSIKA simulation is rather scarce, we rely on n-fold *stratified*[7] *cross-validation* [159] to separate the available simulation into training and validation datasets. The dataset is split into n disjoint subsamples. The fit is performed using n-1 of these subsamples as training dataset and the mlogloss evaluated on the n-th subsample, that has been left out. This procedure is repeated n-times and allows to use the entire simulation for evaluation purposes while maintaining strict separation between training and validation datasets. The result, when averaged over all n validation-folds, is called *out-of-sample (oos)* mlogloss. For model selection purposes $n = 10$ is recommended [163][159][8]. Once the best hyper-parameters (model) are selected, the final fit is performed using the full dataset consisting of all n-folds. Finally, in order to obtain an unbiased estimate of the behavior of this final best-fit model, i.e. background rejection, signal efficiency, passing rates etc., the corresponding properties are studied in a last step using entirely independent *test-datasets* - simulations that have been generated later and never used during optimization. In other words, the final statistical data analysis presented in Chapter 6 only uses simulations that were hidden (unavailable) during parameter optimization and fitting of the GBDT model.

[7] Stratified means that we require the relative contribution from each class (and the various types of CORSIKA simulations with different injection spectra) to each fold to be identical to the entire sample. This ensures that, for the purpose of measuring mlogloss, the n-folds behave identically up to statistical fluctuations.

[8] For reasons of numerical efficiency, we performed most calculations for $n = 2$ and used $n = 10$ only for a couple of points with highest performance.

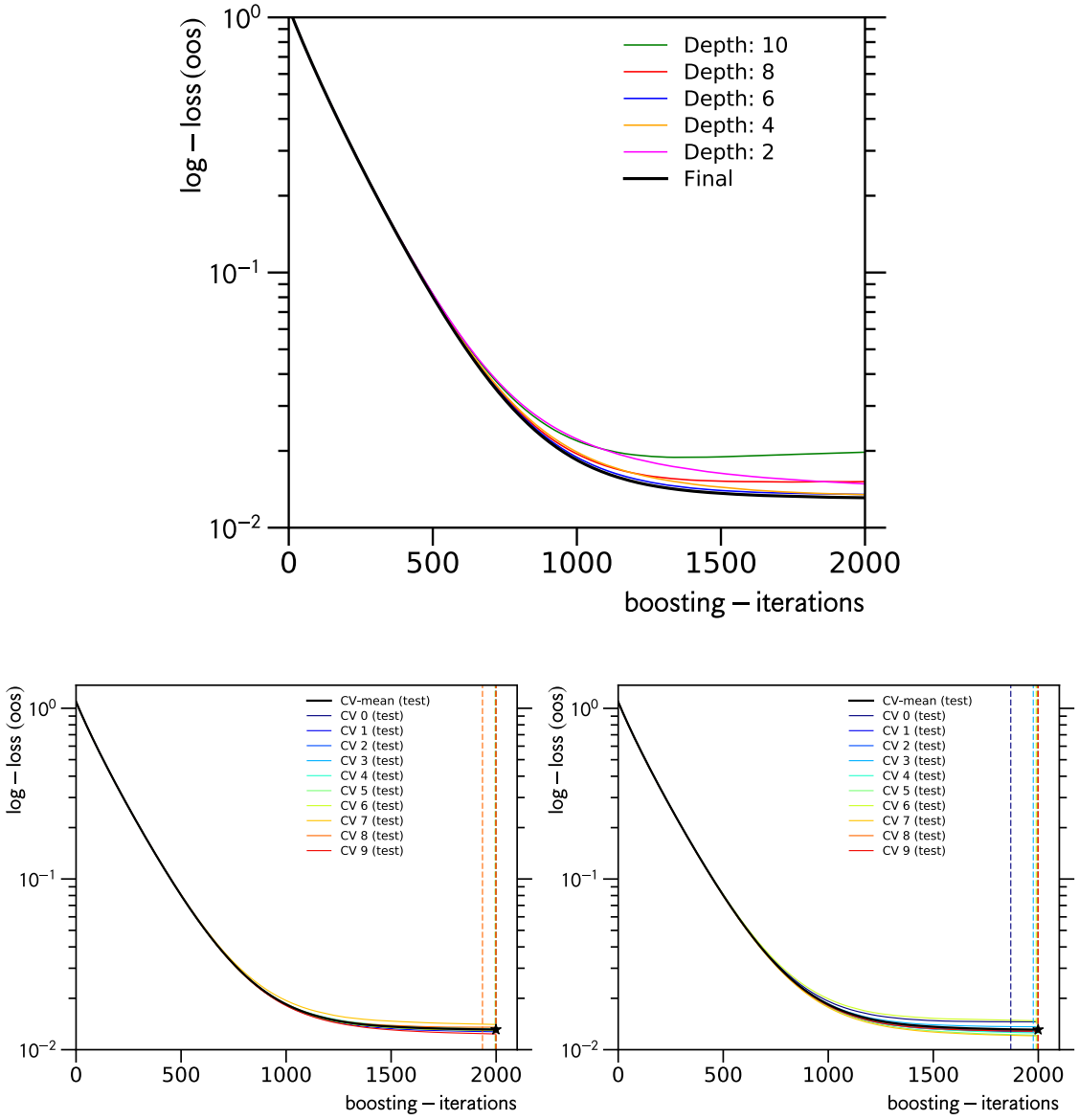


FIGURE 3.12: Optimization of GBDT parameters: Dependence of log-loss criterion on maximum depth of each tree (top). Two examples for log-loss estimation via 10-fold cross-validation for final GBDT parameters, see text (bottom)

Since strong shrinkage (small learning rate η, c.f. eq. (3.26)) has been found to always improve stability [158], we chose a conservatively small value $\eta = 5 \cdot 10^{-3}$ without further optimization. As shown in Fig. 3.12 (top), depending on the maximum depth of the individual trees (varied between $T = 2$ and $T = 10$, c.f. eq. (3.27)) we observe convergence of the out-of-sample mlogloss between $M = 1000$ and $M = 2000$ boosting iterations, where the mlogloss criterion attains its minimum value.. No strong dependence on the tree depth is observed and values between $4 - 6$ appear to work well, except for $T = 10$, which performs worse, indicating overfitting. After some further tests, we observe the best performance for the final model with $T = 8$ and extra regularization (black line). To confirm that this is not an artifact from one particular choice of cross-validation splits,

we have repeated the tenfold cross-validation several times (random splits) and obtained essentially indistinguishable results. Two examples are shown in Fig. 3.12 (bottom left and bottom right). However, since the observed differences in oos-mlogloss are small, we could have probably also chosen any of the two well performing un-regularized models with $T \in \{4, 6\}$ for further study. The values for the hyper-parameters corresponding to the final model are given in Tab. 3.7.
The solution is found using XGBoost [152] and yields a parametrization of the class membership probabilities $p_k(\mathbf{X})$ (eq. (3.18)) for each class k in the problem. For each event with observables $\mathbf{X}$ we can now evaluate these probabilities (c.f. eq. eq. (3.16)): $p(\text{cascade} \,|\, \mathbf{X})$ (cascade score), $p(\text{starting track} \,|\, \mathbf{X})$ (starting track score) and $p(\text{atm. muon} \,|\, \mathbf{X})$ (atm. muon score).

Parameter	Symbol	Value
shrinkage	η	$5 \cdot 10^{-3}$
# of boosting iterations	M	2000
max. depth of individual trees	T	8
regularization of # of terminal nodes	γ	5
regularization of leaf values	λ	1 (default)
min. sum of weights per node	w_{tot}	100

TABLE 3.7: Optimized values for GBDT parameters used in this work.

3.2.3 Level 5A: Extra Cleaning

Using the GBDT model, constructed in Sec. 3.2.2, eq. (3.16), we are now in a position to study background events that appear reasonably cascade-like, i.e. the near-signal region. Of special interest to us are background events that contribute to the tail of mis-reconstructed high energy events $E_{rec} > 10\,\text{TeV}$ and remain after the Level 4 pre-selection (c.f. Fig. 3.6, bottom). Fig. 3.13 (top left) shows that, according to our CORSIKA simulation, almost all these events deposit very little light ($Q_{tot} < 100\,\text{p.e.}$) and thus appear very dim inside the detector. Clearly, the reconstruction completely failed for such events. While their contribution to the total number of low-charge ($Q_{tot} < 100\,\text{p.e.}$) events is negligible ($< 0.1\%$), they pose a problem. The total rate of such low-energy background events is sufficiently large, that even this tiny fraction of mis-reconstructed events would dominate the expected signal of high-energy $E > 10\,\text{TeV}$ astrophysical neutrinos. How does the GBDT model respond to such events? Fig. 3.13 (2nd row, left) shows the subsample of all simulated events (CORSIKA) that are classified loosely as somewhat cascade-like with $p(\text{cascade} \,|\, \mathbf{X}) > 0.1$. Most of these events appear easily identifiable as atmospheric muon background and the total background contribution is strongly reduced. However several (~ 5) simulated CORSIKA events survive in the critical region of low-charge deposits and (mis-reconstructed) high energy. We studied the events in detail. While very dim, they do not show signs of track-like light emission and thus the GBDT probability assignment appears reasonable. However all of these events are located approximately at the same depth ($-135\text{m} < \text{Monopod.z} < -85\,\text{m}$), where the dust-layer (cf. Sec. 2.2) shows peak dust concentrations and hence exceptionally large absorption. Thus such events are easily removable. However due to their small number,

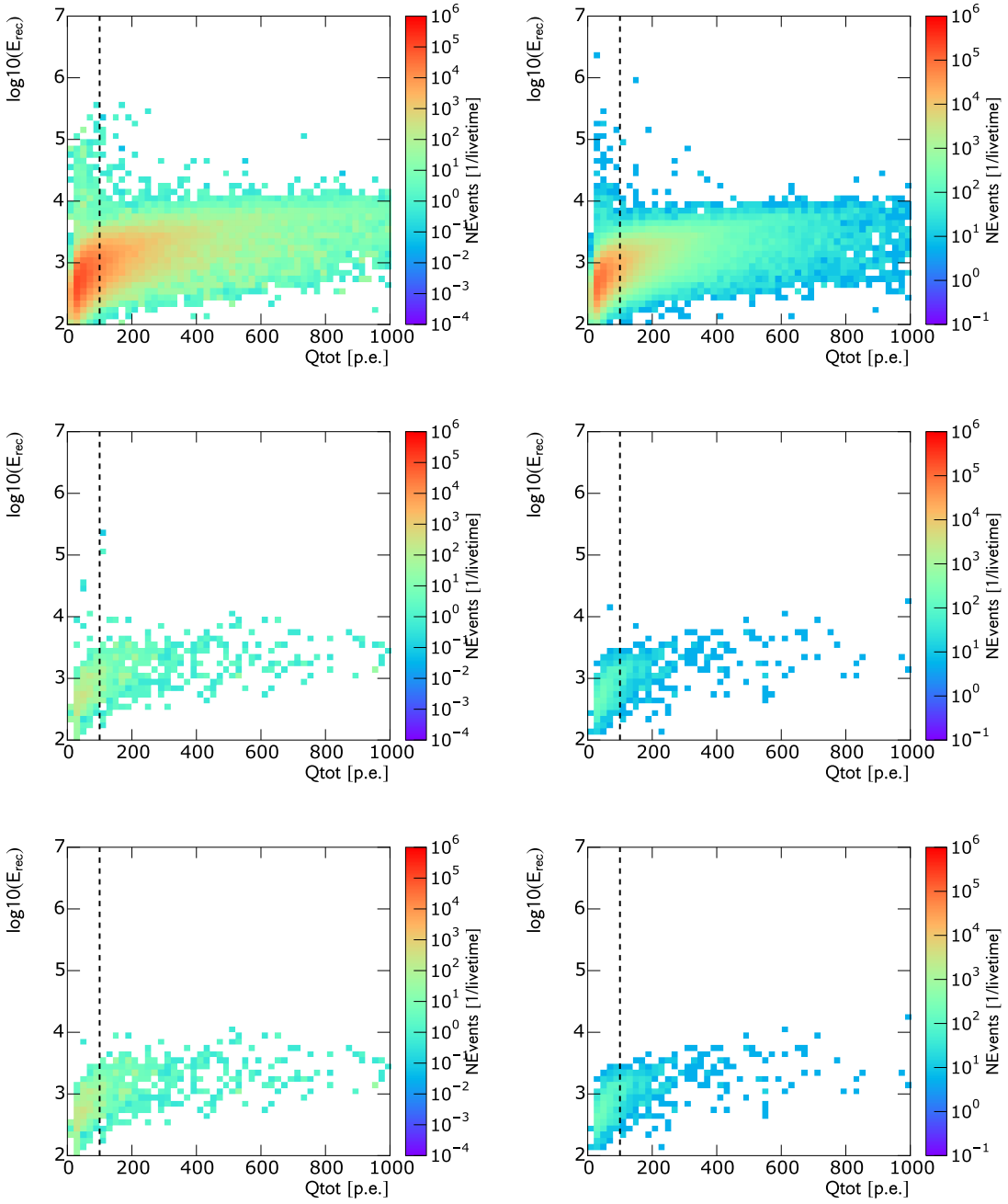


FIGURE 3.13: Population of low-charge miss-reconstructed events: Two dimensional distributions of reconstructed energy and total charge. CORSIKA simulation (left) and data (right). all events after Level-4D (top). events after Level-4D with cascade.score > 0.1 (2nd row). events after Level-5A with cascade.score > 0.1 (bottom). see text for details.

it is difficult to devise a robust cut. We will show later, that such low charge, low (true) energy events pose another challenge in general: due to their large rate in combination with strong background rejection, enormous amounts of background simulations would

be needed to accurately characterize the corresponding residual contribution that would remain once we require $p\,(\text{cascade}\,|\,\mathbf{X})$ to be large. Since the contribution from astrophysical neutrinos (signal) to such dim events is negligible compared to all atmospheric backgrounds, we decided to remove all events with low charge $Q_{tot} < 100\,\text{p.e.}$. Finally, to further stabilize the energy reconstruction, we require at least 4 strings[9] to be active, i.e. to have registered PMT hits in hard local coincidence (c.f. Sec. 2.3), for events with reconstructed energies above 3 TeV. The selection criteria read

[9]The Cascade-L3 filter, starting point of this analysis, requires at least 3 strings to be active.

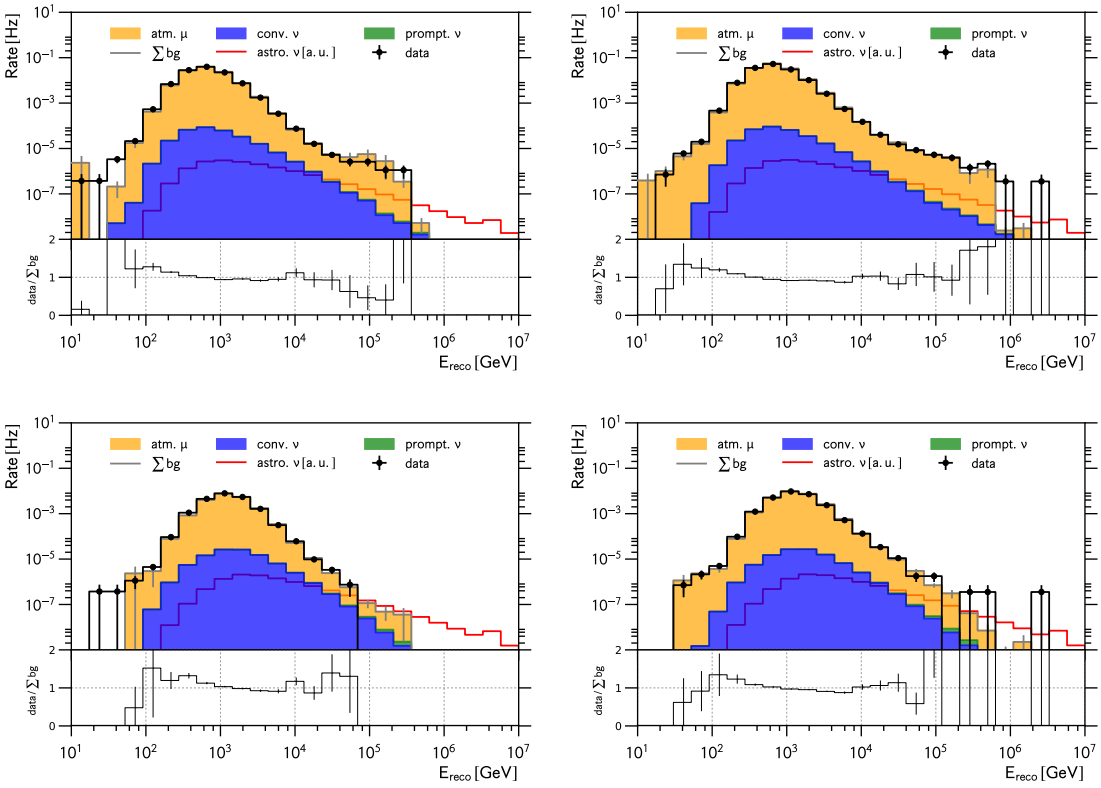


FIGURE 3.14: Before/After comparison of reconstructed energy distributions - Top: Reconstructed energy for 2011 (left) and 2012 (right) configuration after Level-4D selections; Bottom: Reconstructed energy for 2011 (left) and 2012 (right) configuration after Level-5A selection.

Rate in Year	**2015** [mHz]	**2014** [mHz]	**2013** [mHz]	**2012** [mHz]	2011 [mHz]
Data (10%)	$\mathbf{23.18 \pm 0.09}$	$\mathbf{23.25 \pm 0.09}$	$\mathbf{23.33 \pm 0.09}$	$\mathbf{26.3 \pm 0.1}$	21.02 ± 0.09
CORSIKA	n/a	n/a	n/a	$\mathbf{26.9 \pm 0.1}$	20.1 ± 0.2

TABLE 3.8: After Level-5A selection: Passing rates of experimental data for data taking periods 2011-2015 in comparison to estimated passing rates from Monte Carlo simulations (CORSIKA). n/a ≡ simulation not available.

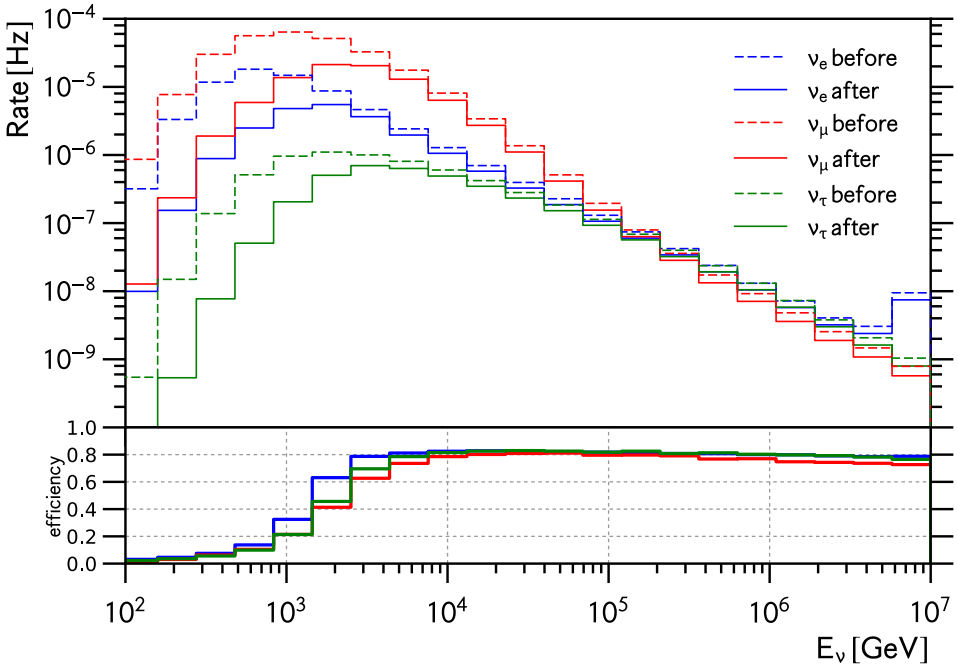


FIGURE 3.15: Top: Neutrino selection efficiencies of all Level-4 + Level-5A criteria with respect to Cascade-Level3 (SC) filter

Keep event if

$$\begin{aligned} & Q_{tot} > 100\,\text{p.e.} \\ \text{and} \quad & (\text{Monopod.z} < -135\,\text{m} \quad \text{or} \quad \text{Monopod.z} > -85\,\text{m}) \\ \text{and} \quad & (\text{NString} \geq 4 \quad \text{or} \quad \text{Monopod.energy} < 3\,\text{TeV}) \end{aligned} \tag{3.29}$$

otherwise reject it.

Fig. 3.14 compares the energy spectra obtained after application of this Level 5A cleaning filter (bottom) to the one after application of all Level 4 selections (top) for the two test-datasets (left: 2011, right: 2012). The reduction of mis-reconstructed muon background at high energies is very noticeable. The Level 5A passing rates are given in Tab. 3.8. Compared to Level 4 the total data volume is reduced by $\sim$ 80%, with the reduction being dominated by the $Q_{tot} > 100\,\text{p.e.}$ requirement. As shown in Fig. 3.15 the neutrino selection efficiency compared to Cascade Level-3 for neutrinos with energies above a few TeV is essentially similar to that of our Level 4 selection $\sim$ 80%. However the $Q_{tot} > 100\,\text{p.e.}$ cut depletes the sample of very low energy neutrinos and thus the efficiencies are low: $< 30\%$ for $E_\nu < 1\,\text{TeV}$, as expected. The expected neutrino passing rates are given in Tab. 3.9 and remain a factor of $\sim$ 250 smaller than the atmospheric muon background expectation.

Rate in Year	**2015** [μHz]	**2014** [μHz]	**2013** [μHz]	**2012** [μHz]
ν_e (conv.)	c.f. 2013	c.f. 2013	$\mathbf{16.47 \pm 0.03}$	$\mathbf{16.47 \pm 0.03}$
ν_μ (conv.)	c.f. 2013	c.f. 2013	$\mathbf{85.4 \pm 0.4}$	$\mathbf{85.5 \pm 0.4}$
ν_e (prompt)	c.f. 2013	c.f. 2013	$\mathbf{0.5388 \pm 0.0005}$	$\mathbf{0.5390 \pm 0.0005}$
ν_μ (prompt)	c.f. 2013	c.f. 2013	$\mathbf{0.1587 \pm 0.0005}$	$\mathbf{0.1592 \pm 0.0005}$
ν_e (astro)	c.f. 2013	c.f. 2013	$\mathbf{3.526 \pm 0.003}$	$\mathbf{3.528 \pm 0.003}$
ν_μ (astro)	c.f. 2013	c.f. 2013	$\mathbf{1.274 \pm 0.003}$	$\mathbf{1.286 \pm 0.003}$
ν_τ (astro)	c.f. 2013	c.f. 2013	$\mathbf{2.579 \pm 0.004}$	$\mathbf{2.580 \pm 0.004}$

TABLE 3.9: After Level-5A selection: Estimated neutrino passing rates for data taking periods 2012-2015 based on NuGen simulation for 2013 configuration

3.2.4 Response to Atmospheric Muon Background

In order to devise a final event selection based on the GBDT classification of each event (multinomial class probabilities), we need an accurate prediction of the background contamination as a function of the classification score towards high scores, for example $p\,(\text{cascade}\,|\,\mathbf{X}) \rightarrow 1$. First, it is necessary to demonstrate that the GBDT classification does not dependent on the details of the detector configuration throughout the years of interest (2012-2015). Fig. 3.16 shows the distributions of the different classification scores ($p\,(\text{cascade}\,|\,\mathbf{X})$ (top), $p\,(\text{starting track}\,|\,\mathbf{X})$ (bottom left) and $p\,(\text{track}\,|\,\mathbf{X})$ (bottom right)) observed in 100% of the experimental data for the different years (colors) - after all pre-cuts (Level 3 + L4 A-D + L5A). The distributions match very well, except for very background like events (e.g. $p\,(\text{cascade}\,|\,\mathbf{X}) \rightarrow 0$) where the differences in the cascade-filter (Level-3) are visible. Such events are of no interest to this analysis and will be removed later. We will now study the cascade-like muon background ($p\,(\text{cascade}\,|\,\mathbf{X}) > 0.1$) in more detail, using the available CORSIKA simulations. Fig. 3.17 compares the background prediction from CORSIKA simulation to the experimental data (left: 2013-2015 experimental data and 2011 CORSIKA[10], right: 2012 experimental data and 2012 CORSIKA). Clearly, the available CORSIKA statistics is insufficient to properly characterize the distribution of the muon background and we observe large statistical fluctuations in the background simulation, dominated by individual events with very high importance weights (small CORSIKA livetime). However we can use the available simulation to establish common properties of simulated background events that appear reasonably cascade-like ($p\,(\text{cascade}\,|\,\mathbf{X}) > 0.1$). Of particular interest is the fraction of muon background events that arrive at the detector (i.e. after passing through the ice shield) as single muons for two reasons. First, the modeling of muon bundles is associated with larger systematic uncertainties than that of single muons and, second, single muon background can more efficiently be generated by means of simulation since it does not require simulation of the full air-shower development (c.f. Sec. 2.5). Fig. 3.18 shows the single muon fraction as estimated from CORSIKA simulation as function of $\log_{10}\left(p\,(\text{cascade}\,|\,\mathbf{X})\right)$ for the two datasets, discussed so far (2011 (top left) and 2012 (top right)). However robust estimation of the single muon fraction remains difficult, especially at larger reconstructed energies, because of limited CORSIKA simulation. For example we only find 8 (22) events in the 2012 (2011) CORSIKA dataset that satisfy

[10]Since 2011 CORSIKA was also used for creation of the GBDT model, we used 20-fold CV to estimate this distribution

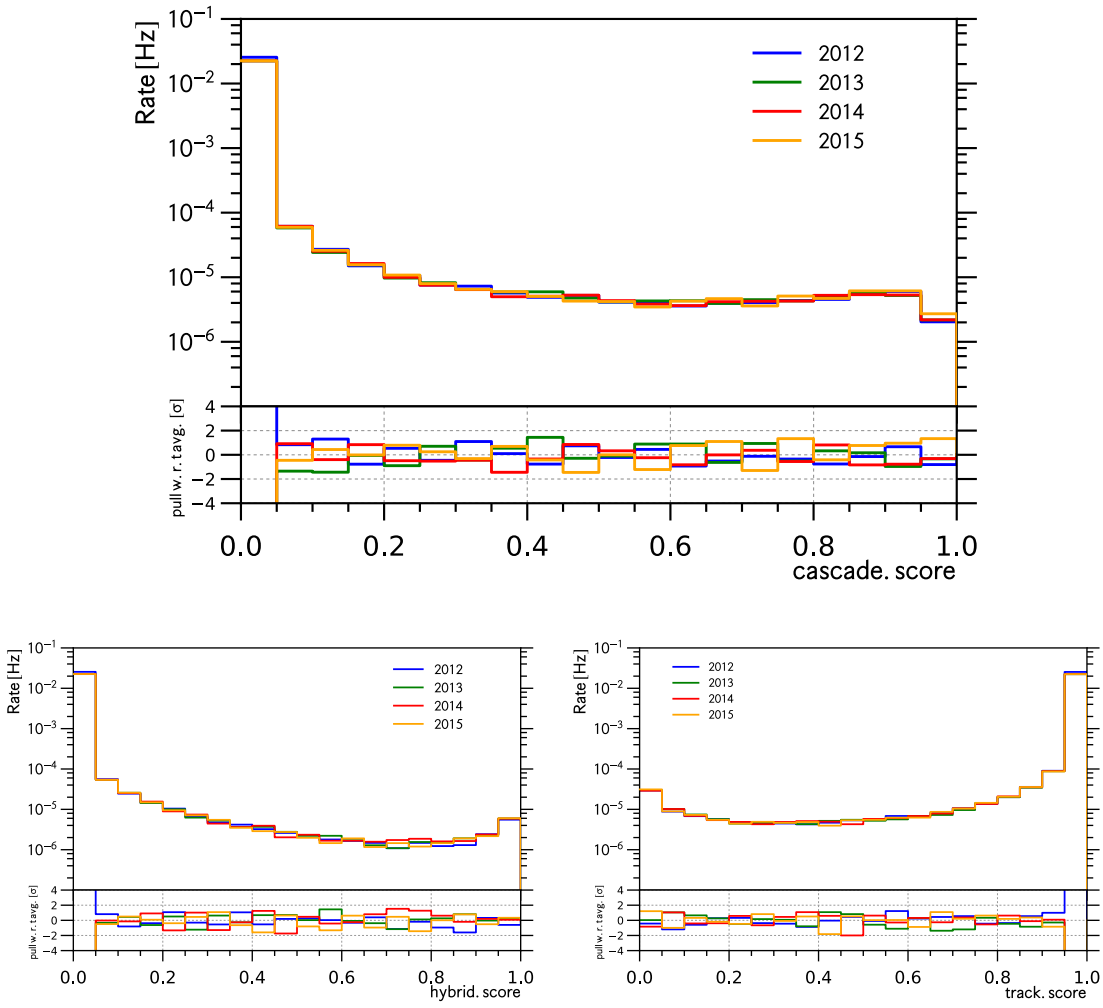


FIGURE 3.16: Distribution of GBDT scores observed in experimental data (100%) for each year (2012-2015). Top: cascade score, Bottom left: starting track score, Bottom right: track score

$E_{rec} > 5\,\text{TeV}$ and $p\,(\text{cascade}\,|\,\mathbf{X}) > 0.1$. Therefore we analyzed yet another CORSIKA dataset (2010) from an older, partial detector configuration (79 strings instead of 86 strings), corresponding to the largest CORSIKA Monte Carlo Simulation production in IceCube. For comparison: in the 2010 CORSIKA dataset we find 58 muon background events with the criteria mentioned above. The corresponding single muon fraction is shown in Fig. 3.18 (center). Some differences between the different CORSIKA datasets (2010/11/12) are visible and are due to differences in the raw (unweighted) cosmic-ray injection spectra and composition. However they all demonstrate that the single muon fraction increases with $p\,(\text{cascade}\,|\,\mathbf{X})$ and thus muon bundles are efficiently rejected. In particular we find $N_{single\ \mu}/N_{tot} \gtrapprox 90\%$ for $p\,(\text{cascade}\,|\,\mathbf{X}) > 0.1$. This justifies to estimate the behavior of the atmospheric muon background in this region using fast single muon simulation ("MuonGun" [139], c.f. Sec. 2.5). For comparison, the currently available MuonGun simulation contributes $\sim$ 1400 events to the region with $E_{rec} > 5\,\text{TeV}$ and $p\,(\text{cascade}\,|\,\mathbf{X}) > 0.1$. Fig. 3.18 (bottom) demonstrates an interesting behavior of

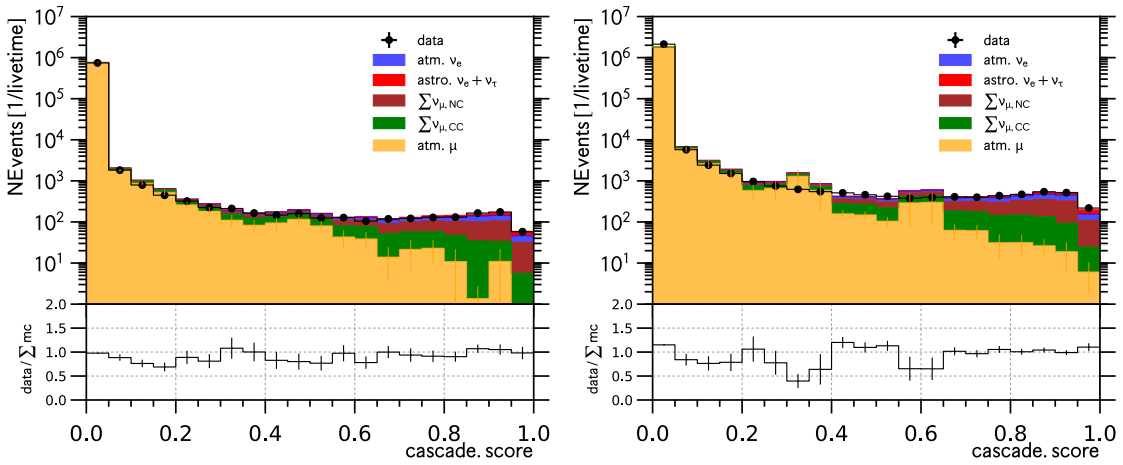


FIGURE 3.17: Comparison of cascade.score distributions observed in data to expectation from simulation. Left: data (100%), simulation (CORSIKA + NuGen) for season 2012. Right: data (100%) for seasons 2013-15 , NuGen simulation for season 2013 and CORSIKA simulation for season 2011. Notice the large MC statistical uncertainties on the prediction from CORSIKA (orange) at high cascade.scores.

single muons as function of $p\,(\text{cascade}\,|\,\mathbf{X})$: the fraction of the largest, single radiative muon energy loss (e.g. bremsstrahlung and pair-production, c.f. Sec. 1.8) in the detector compared to the total muon energy at detector entry increases with the cascade.score[11] . This is exactly what one would expect from our GBDT model. Single energy depositions appear point-like in the detector and thus are indistinguishable from the signature of neutrino-induced cascades. Hence, in order to be recognizable as muon track, the muon needs to deposit sufficiently large visible energy over some extended range in the detector. However if the muon deposits most of its energy in a single cascade-like energy loss, then it might not have sufficient energy left to be detectable. Fig. 3.18 (bottom) shows that 50% (80%) of the $E_{rec} > 5\,\text{TeV}$ muons that contribute at the highest cascade probabilities deposit more than 96% (93%) of their remaining energy at detector entry in a single cascade-like radiative energy loss. In conclusion the main muon background that contributes at high cascade scores is a tail of highly stochastic single muons.

Using - from now on - single muon simulation dramatically increases the available simulation statistics. Fig. 3.19 compares the distributions of the different classification scores (cascade score (top), starting track score (botton left) and muon track score (bottom right)) obtained from single muon (and neutrino) simulation to the full experimental dataset (2012-2015). Data and simulation agree very well, except for cascade and starting track scores $\to 0$ (track.score $\to 1$), where muon bundles (not simulated) contribute significantly, as discussed above. The simulation predicts a smooth decrease of single muon background towards large signal scores (neutrino induced cascades and starting tracks). Similarly, at large starting track scores the neutrino simulation predicts an essentially pure contribution from ν_μ charged-current interactions, while at large cascade

[11]This is the energy that the muon has left after having passed through the ice-shield - not the muon energy at the surface of the Earth.

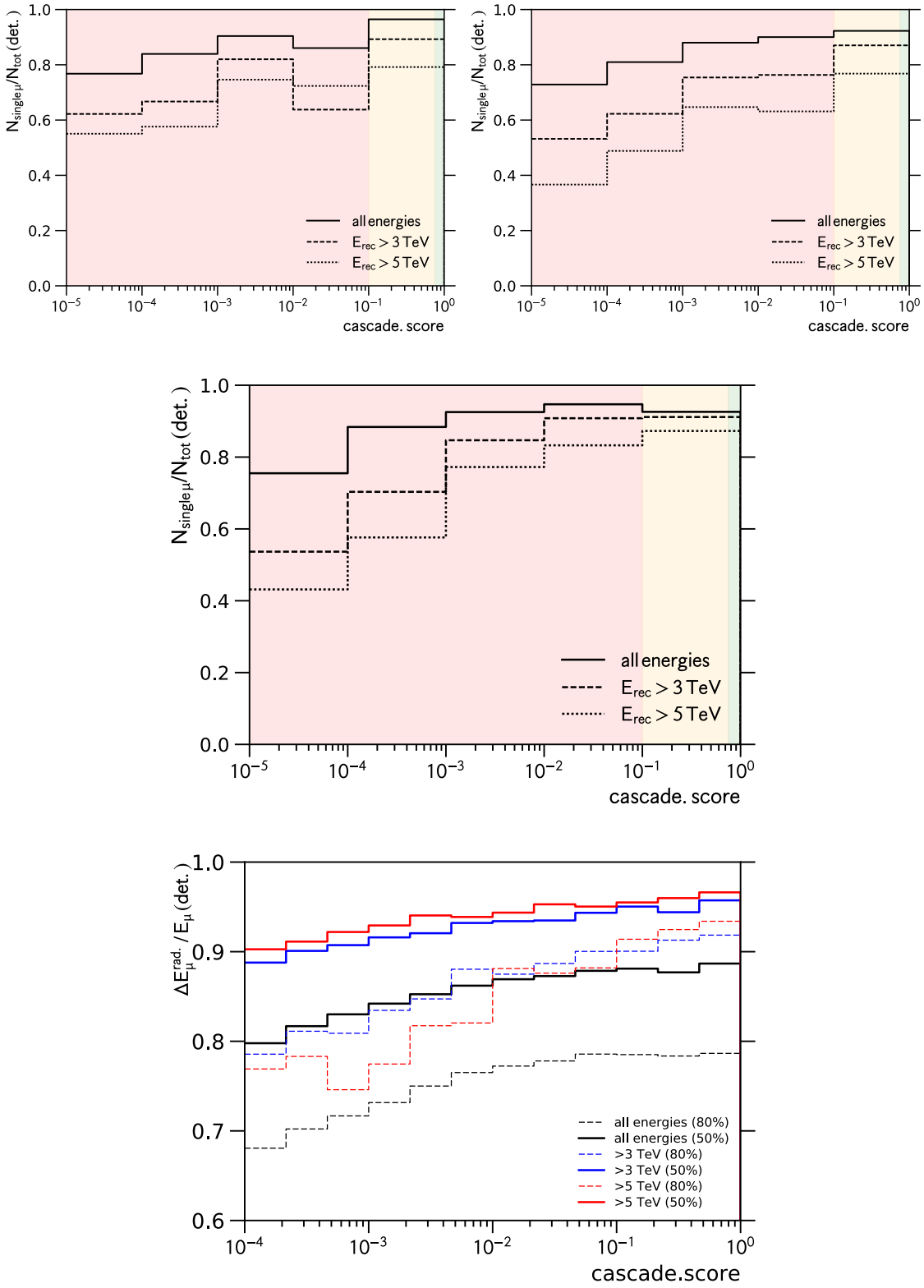


FIGURE 3.18: Fraction of muon background events that arrive at the detector as single muons as function of cascade.score for different CORSIKA datasets: CORSIKA 2011 (top left), CORSIKA 2012 (top right), CORSIKA (center). Bottom: Fraction of largest radiative energy loss in the detector relative to total muon energy at detector entry as function of cascade.score - estimated from MuonGun simulation.

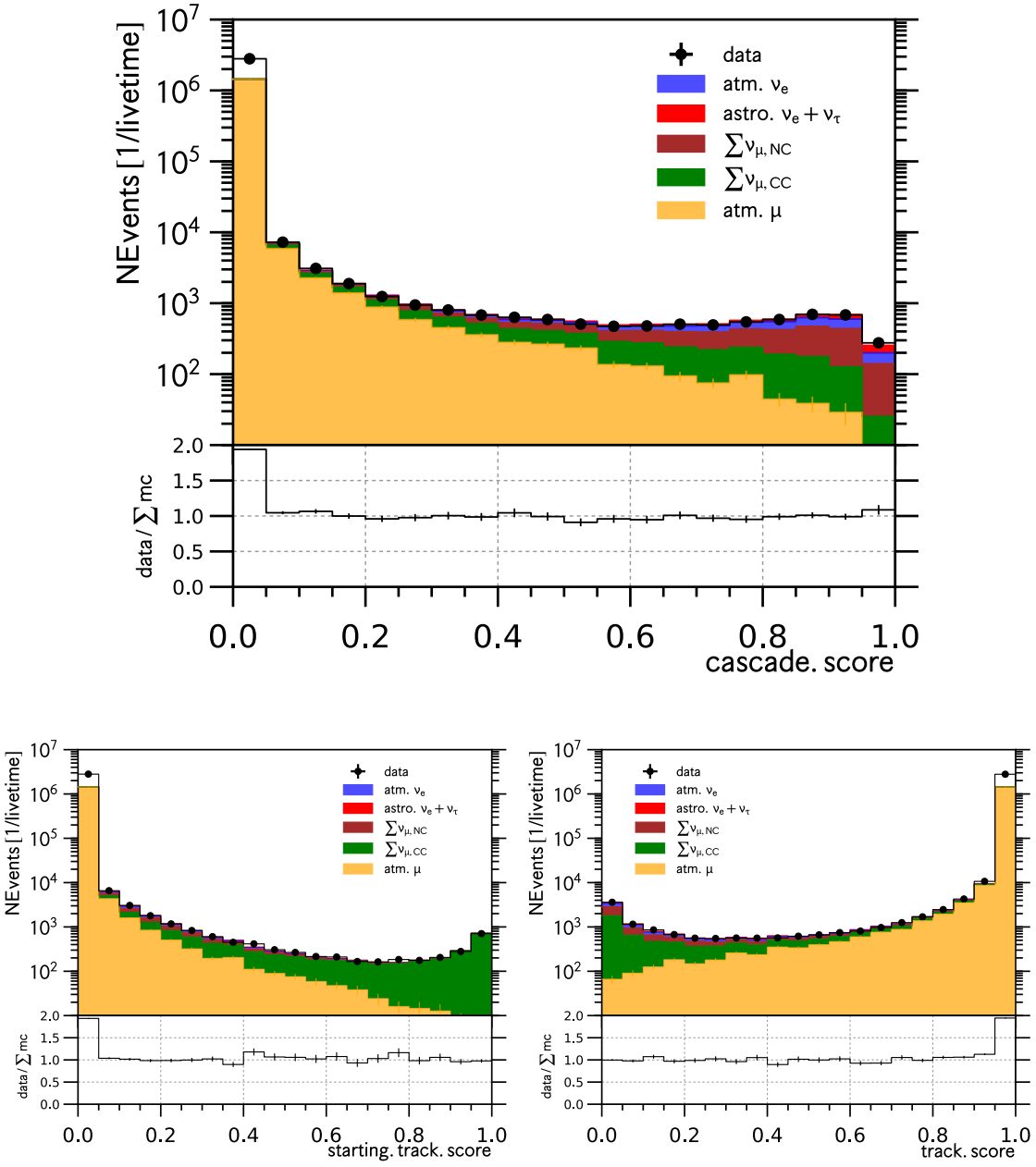


FIGURE 3.19: GBDT Score Distributions: data (100%) for 2012-2015 and simulation (MuonGun + NuGen) for 2013: cascade score (top), starting track score (bottom left) and track score (bottom right).

scores neutrino induced cascades dominate (ν_e CC interactions + NC interactions from all flavors).

Since probabilities must sum to unity, the three classification scores are correlated and the classification can be visualized in two dimensions. Fig. 3.20 shows the joint distribution of all three classification scores as ternary projection for the full experimental dataset (2012-2015). The three corners in the triangle represent regions of high purity of each class:

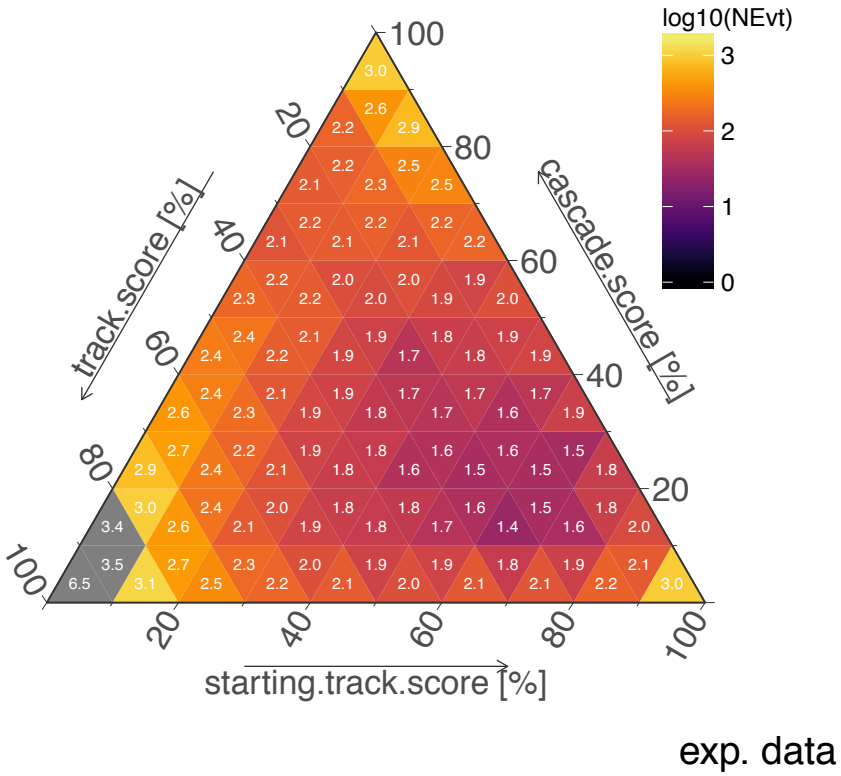


FIGURE 3.20: Ternary projections of GBDT classification for experimental data

atmospheric muon tracks (bottom left), cascades (top) and starting tracks (bottom right). The number of events in each triangular bin compare well to the simulation prediction, which is shown in Fig. 3.21 (top left) and includes single muon background as well as neutrino simulation. Thus, the correlations between the classification scores are modeled correctly by our simulations. The decomposition of the simulation prediction into the different signatures/contributions are also given: cascades, i.e. ν_e+ν_τ+ν_μ-NC (2nd row, right), starting tracks, i.e. ν_μ-CC (bottom left) and single muons (bottom right) and clearly illustrates that the GBDT model successfully discrimination between the different event classes. We will use all three scores to define the final selection, consisting of a cascade signal sample (c.f. Sec. 3.2.6), a ν_μ-CC control sample (c.f. Sec. 3.2.7) and a single muon control sample (c.f. Sec. 3.2.8).

3.2.5 Level 5: GBDT Selection Criteria

When designing selection criteria, several points have to be considered: purity, sample size and systematics. In a discovery experiment, one might optimize the selection based on the expected significance given some assumed value for the unknown strength of the signal. In a measurement like ours, where the discovery itself is not of primary interest anymore, since the existence of the signal has already been established, one might instead

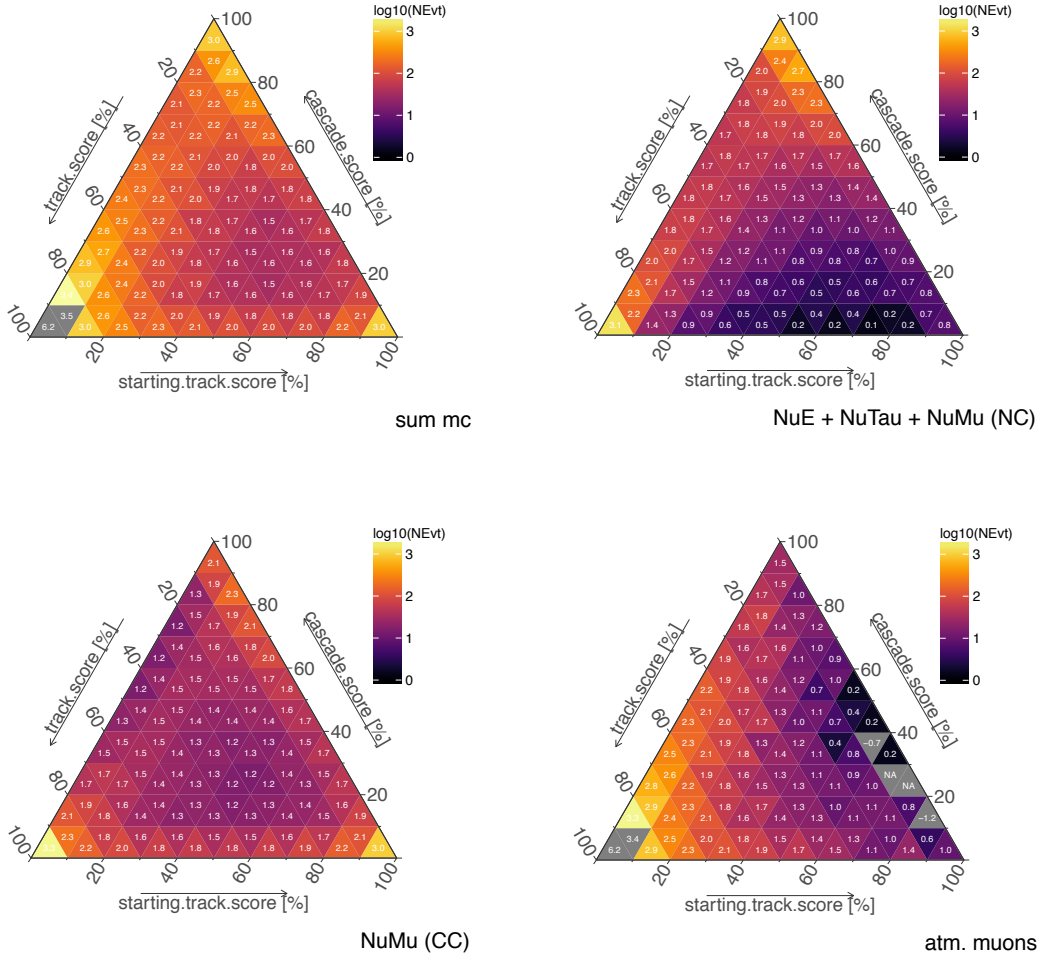


FIGURE 3.21: Ternary projections of GBDT classification for simulation: sum of MC (top left), sim. cascade-like events (top right), sim. starting track events (bottom left) and sim. single muon events (bottom right)

choose the cuts such that they minimize the expected uncertainty in the measurement of the model parameters. In the context of this work this would mean to minimize the expected uncertainties for the parameters of the astrophysical neutrino flux. Since we anticipated systematic uncertainties, for example the detector response to light signals, (c.f. Sec. 4) to be non-negligible, especially at low energies, these uncertainties would need to be taken into account during such optimization. However during development of the selection criteria insufficient neutrino Monte Carlo simulations corresponding to alternative detector properties (e.g. variations in optical properties of the ice) were available, thus rendering "blind" optimization of the selection, based on a single metric, impossible. Previous IceCube masurements [164][155] showed that neutrino flux measurements work well if the atmospheric muon contribution is suppressed to the $\sim 10\%$ level. Thus requiring $< 10\%$ muon background became a first desideratum. We then aimed to maximize the cascade acceptance of the selection, subject to the aforementioned constraint. Finally, we need to separate starting tracks (dominated by conventional atmospheric muon neutrino

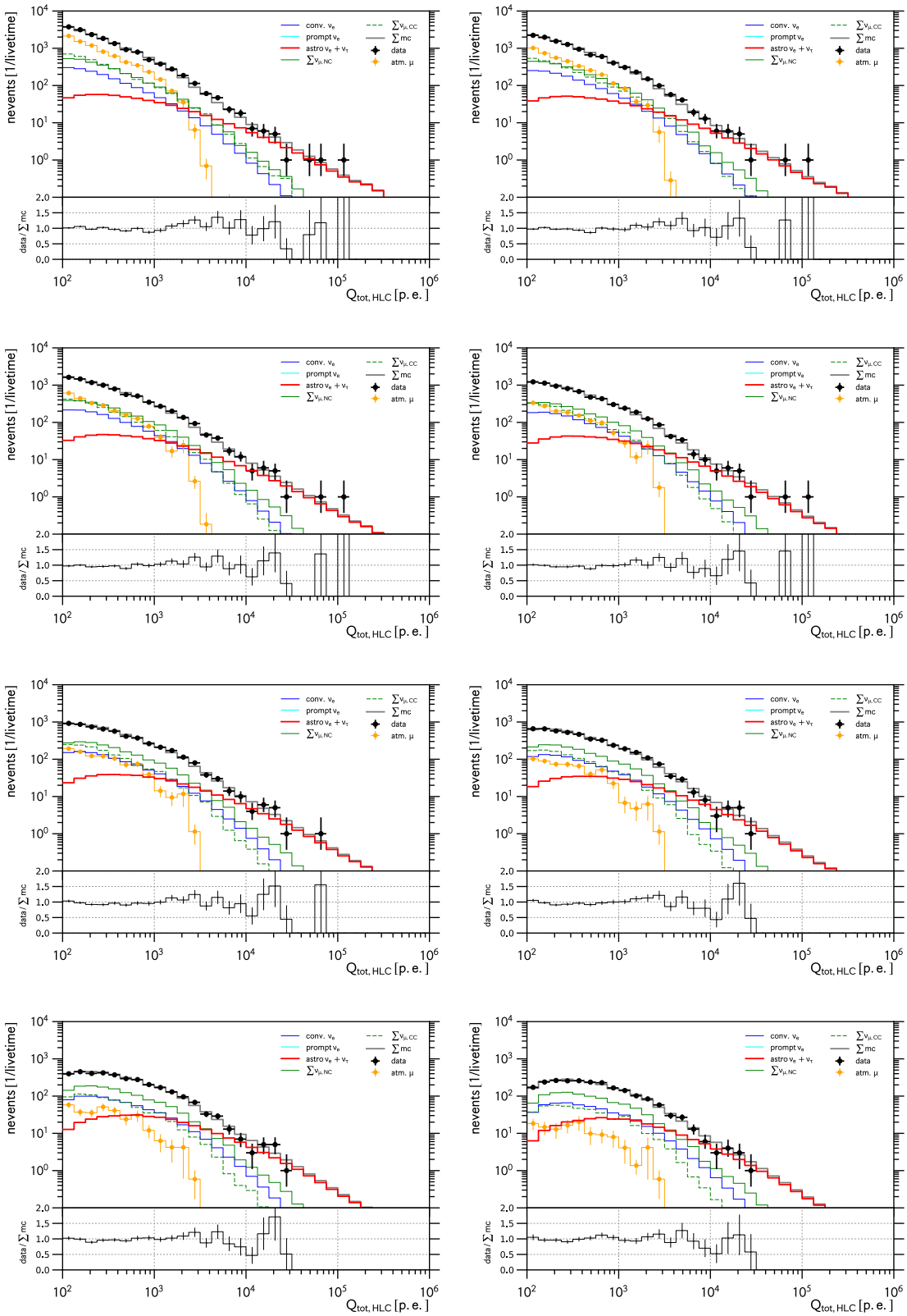


FIGURE 3.22: Observed distribution of total PMT charge: 2012-2015 data (100%) and simulation (MuonGun + NuGen). Different cascade.score thresholds in steps of 0.1: from > 0.1 (top left) to > 0.8 (bottom right)

background) from cascades to control for uncertainties in the conventional atmospheric neutrino flux. Since starting tracks have significantly larger atmospheric neutrino background than cascades, mis-classification of cascades as starting tracks is considered worse than mis-classification of starting tracks as cascades. Being unable to study the robustness of the starting track classification against systematic uncertainties, we decided to choose a conservative cut, corresponding to high starting track purity in the ν_μ-CC control sample. As shown in Fig. 3.19 (bottom left) a sufficiently large and pure sample of starting tracks can be isolated by requiring a large value for the starting track score. We studied many simulated ν_μ-CC events (and eventsin the 10% exp. test data) from that region in the event viewer. The starting muon produced in ν_μ-CC interaction appeared well identified for events with hybrid.score > 0.75. Hence the first selection criterion becomes

Add event to Starting Track Sample if

$$\text{starting.track.score} \geq 0.75 \tag{3.30}$$

otherwise reject it from Starting Track Sample.

This cut corresponds to a straight line at starting.track.score $> 75\%$ parallel to the left edge of the ternary projection (Fig. 3.20) thus isolating the bottom right corner. The properties of this sample are discussed in more detail in Sec. 3.2.7.

Since the measurement of the astrophysical neutrino flux will rely mostly on information from the cascade signal sample, we studied the behavior of the cascade.score variable in more detail. In particular, we want to make sure that the agreement between data and simulation does not depend strongly on the exact value of the cascade.score cut. That this is the case is demonstrated in Fig. 3.22 using the total charge Q_{tot}^{HLC} as a measure of the brightness of each event. The data (black) compares very well to the simulation prediction for the various components: single muons (yellow) and neutrinos (other colors) independent of the cascade.score requirement, which is varied in steps of $\Delta = 0.1$ from cascade.score > 0.1 (top left, single muon dominated) to cascade.score > 0.8 (bottom right, cascade dominated). The figure nicely shows how the single muon background is reduced as a function of the cascade.score. In addition the signal to background ratio improves with total charge for higher cascade.scores, suggesting a dependence on the energy of the events.
Fig. 3.23 shows the cascade.score distribution for different values or reconstructed energy threshold: $> 1\,\text{TeV}$ (top left), $> 5\,\text{TeV}$, $> 10\,\text{TeV}$ and $> 25\,\text{TeV}$ (bottom right). The background rejection clearly improves with higher reconstructed energies, as expected. First, high-energy muons are more easily identified than low-energy ones[12] and second, the muon flux decreases rapidly as function of muon energy. We have shown in Sec. 3.2.4 that the fractional contribution from muon bundles at low cascades.scores cascade.score ≈ 0 increases with energy. This explains why the data excess at cascade.score ≈ 0 over the single muon prediction from "MuonGun" increases with energy.
The improved separation between atmospheric muons and neutrino induced cascades at higher energies motivates a two-dimensional cascade.score requirement. A tight cut at low energies suppresses the majority of the muon background, while at higher energies, a

[12] In order to avoid detection, the muon has to essentially loose all of its energy in a single energy loss.

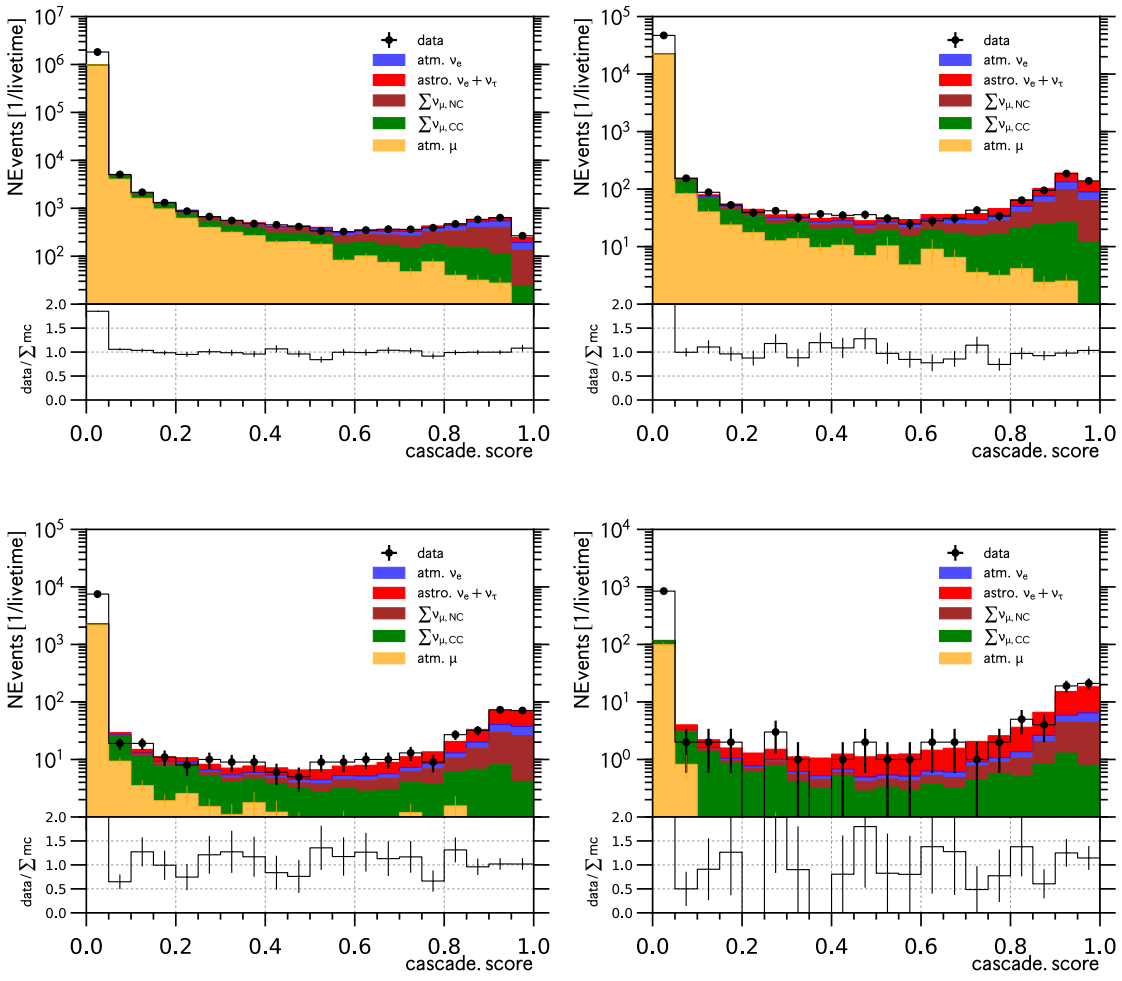


FIGURE 3.23: Observed cascade.score distributions for different energy thresholds: $E_{rec} > 1\,\mathrm{TeV}$ (top left), $E_{rec} > 5\,\mathrm{TeV}$ (top right), $E_{rec} > 10\,\mathrm{TeV}$ (bottom left) and $E_{rec} > 25\,\mathrm{TeV}$ (bottom right).

looser cut can boost signal efficiency without sacrificing background rejection. Fig. 3.24 shows the two-dimensional joint distribution of the cascade.score and reconstructed energy for the entire data sample (2012-2015) (top left), the sum of all neutrino simulations (top right), all available single muon simulation (2nd row, left) and CORSIKA simulation for the years 2010 (2nd row, right), 2011 (bottom left) and 2012 (bottom right).
At low energies $E_{rec} < 5\,\mathrm{TeV}$ a cascade.score > 0.75 is required to keep the atmospheric muon background below the 10% level. At highest energies not one simulated muon background event, neither from single muon simulation nor from CORSIKA full air-shower simulation is rated by the GBDT with cascade.score > 0.1. To smoothly connect these two extreme values we chose to parametrize the energy dependence of the cascade.score cut value as an inverted sigmoid function, defined by the following selection criterion

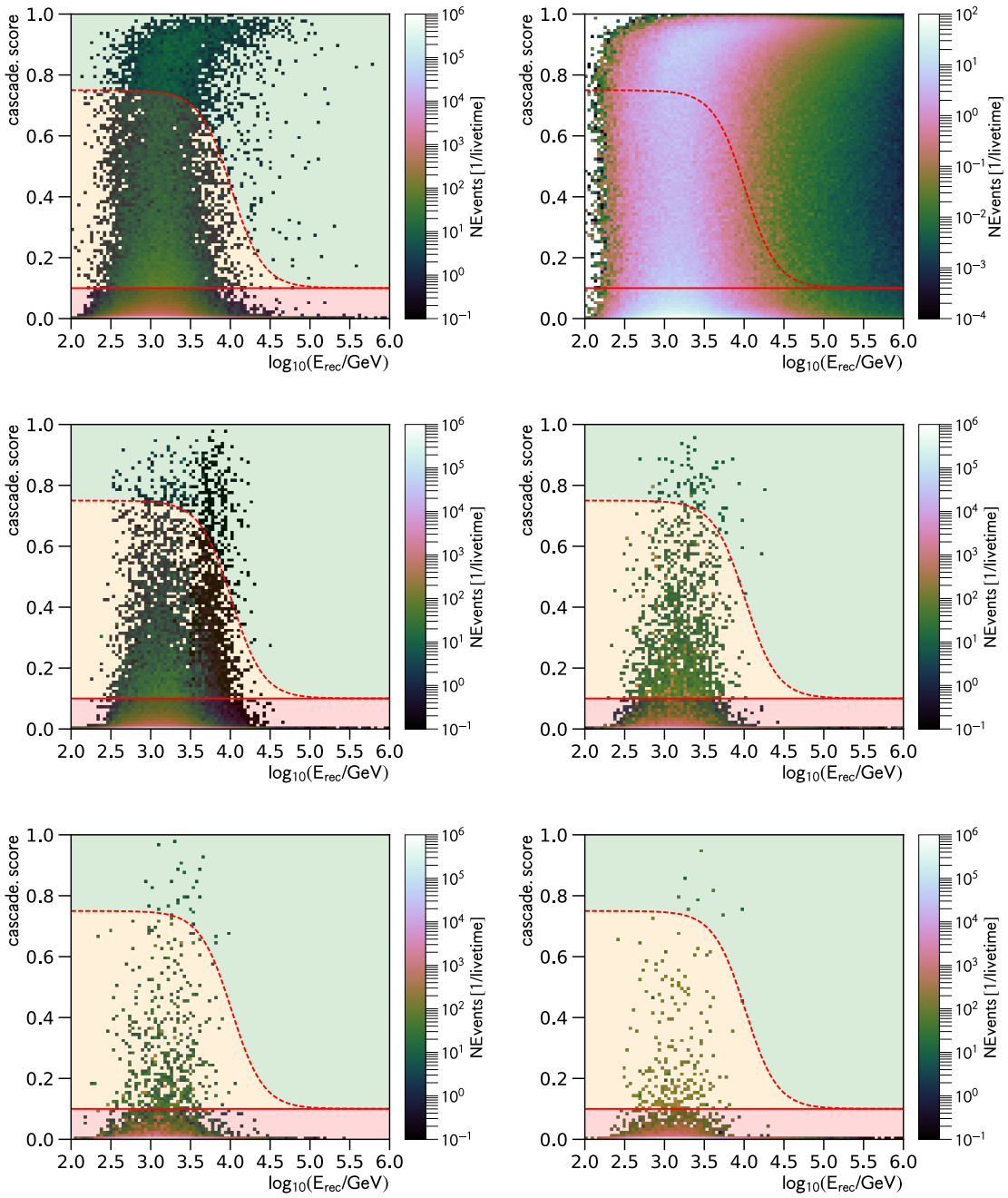


FIGURE 3.24: Two-dimensional distributions of cascade score and reconstructed energy. data (top left), NuGen (top right), MuonGun (center left), Corsika 2010 (center right), Corsika 2011 (bottom left) and Corsika 2012 (bottom right).

Add event to Cascade Signal Sample if

$$\begin{aligned} &\text{cascade.score} > f(\xi), \quad \xi \equiv \log_{10} E_{rec}/GeV \\ &f(\xi) = 1 - \left\{ \frac{1}{A + \exp(-B * (\xi - C))} + D \right\} \\ &A = 1.539,\ B = 5.0,\ C = 4.1,\ D = 0.25 \\ &\text{and} \quad \text{starting.track.score} < 0.75 \end{aligned} \tag{3.31}$$

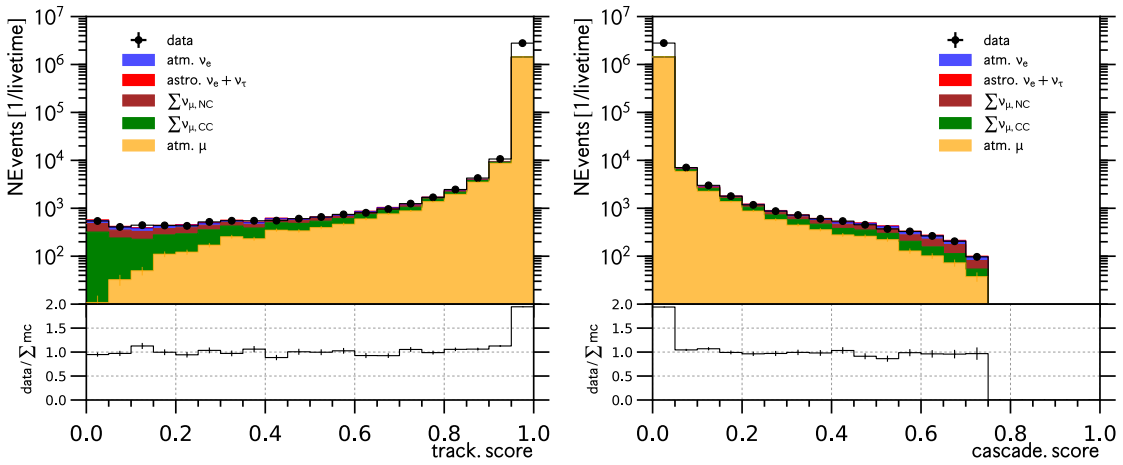


FIGURE 3.25: Track score (left) and Cascade score (right) distributions after application of cuts eq. (3.30) and (3.31), see text.

otherwise check track.score condition, eq. (3.31).

where the last requirement keeps the cascade sample disjoint from the starting track sample. The cut boundary defined by eq. (3.31), is shown in Fig. 3.24 as red, dotted line and the signal region is denoted by the green area. In the ternary projection, Fig. 3.20 this cut runs parallel to the lower edge of the triangle and isolates the top corner (cascades) as a function of energy.

The selection criteria defined so far, eq. (3.30) and (3.31), retain events that are well-identified either as cascades or starting tracks. However the GBDT model provides more information and thus allows to further enhance the neutrino selection efficiency, while maintaining $< 10\%$ muon contamination. Fig. 3.25 (top left) shows the distribution of the track.score variable, after all events that satisfy either eq. (3.30) or (3.31) have been removed. The distribution remains well described by our simulations and neutrinos dominate at low track.score < 0.15. These are events that the GBDT model clearly identifies as neutrinos but does not allow to distinguish further between the neutrino signatures, for example because the starting muon is too low in energy to be detected. Thus we add them to the cascade sample:

Add event to Cascade Signal Sample if

$$\begin{aligned} &\text{track.score} < 0.15 \\ \text{and} \quad &\text{starting.track.score} < 0.75 \end{aligned} \tag{3.32}$$

otherwise reject it from Cascade Signal Sample.

where the last requirement keeps the cascade sample disjoint from the starting track sample. The Cascade sample thus consist of events that satisfy either eq. (3.31) or eq. (3.32).

Finally, using events that are rejected from both neutrino samples (cascades and starting tracks) we would like to define a single muon sample to control the corresponding background in our final analysis. Fig. 3.25 (top right) shows the cascade.score distribution for such events. Good agreement between data and MC simulation is observed, and thus we move all remaining events with cascade.score > 0.1 into the muon control sample.

Add event to Single Muon Sample if

$$\begin{aligned} &\text{event fails eq. (3.30)} \\ &\text{and event fails eq. (3.31)} \\ &\text{and event fails eq. (3.32)} \\ &\text{but event satisfies: cascade.score} > 0.1 \end{aligned} \tag{3.33}$$

otherwise reject event from Single Muon Sample.

Thus, events that do not satisfy at least one of the equations, (3.30) to (3.33), are completely discarded. These events have cascade.score < 0.1 and starting.track.score < 0.75. The former condition is shown in Fig. 3.24 as red, solid line with corresponding rejection region marked as red area. The region between the red solid and red dotted lines defines the single muon control sample and is marked yellow.

3.2.6 Level 5B: The Cascade Signal Sample

The cascade signal sample is defined by eqs. (3.31) and (3.32). We have claimed earlier in this work, that differences in filter configurations, visible at early stages of this selection, e.g. Sec. 3.2.3, vanish in the signal region. Fig. 3.26 (top) shows the energy spectra (left: true neutrino energy, right: reconstructed energy) for the 2012 (dashed lines) and 2013 (solid lines) filter configurations predicted from electron (blue), muon (red) and tau (green) neutrino simulations after the cascade selection criteria are applied. Essentially no difference is observed for electron and tau neutrinos contributing to the cascade sample, while muon neutrinos agree within $< 2\%$ dominated by statistical uncertainties in the simulations[13]. Fig. 3.26 (bottom) demonstrates that the agreement is independent of arrival direction (zenith angle) using electron neutrino simulation as example (left: true variables, right: reconstructed observables). Tab. 3.10 gives the passing rates for the experimental data (2012-2015) and Monte Carlo simulation (2013). In total this sample retains $\sim$ 4700 data events. To better illuminate the properties of these cascade-like events we present relevant observable distributions in Fig. 3.27: total charge (top left), reconstructed energy (top right), cascade.rlogl (2nd row, left), reconstructed zenith angle (2nd row, right), reconstructed vertex z-position (bottom left) and reconstructed vertex XYscale (bottom right). These can be compared to their counterparts after the pre-GBDT cuts (Figs. 3.8 and 3.9 (top)). The GBDT suppresses the data rate strongest in the vicinity of the geometric boundaries defined by the pre-selection criteria (c.f. Sec. 3.1.3 - 3.1.5). Similarly the reduction is strongest in the region of low cascade reconstruction quality (high values of Cascade.rlogl). While our simulation predictions agree well with the observed data in all observables, two observations stand out. First, the simulations

[13]For the final analysis we have since doubled the available muon neutrino simulation.

significantly over-predict the number of events that are reconstructed as vertically upgoing. While still under study, we would just like to point that this direction would maximize the intersection of photon trajectories with the hole-ice column of the nearest string. We will show in Sec. 4.1.2 that the optical properties within the (re-frozen) ice in IceCubes drill holes have large systematic uncertainties and directly impact the directional reconstruction of cascades. Since the relative contribution of events with such reconstructed directions to the entire sample is small, this is not expected to negatively impact this analysis. We will refer to Sec. 6 and Appendix A.2 for further analysis and discussions. Second, the agreement between data and simulation prediction near the upper boundary of the selection ($Monopod.z > 200\,\mathrm{m}$) appears worse than it is throughout the remaining part of the detector. This does not appear to be caused by the GBDT selection, since we have observed a similar trend after all pre-cuts before application of the GBDT, as shown in Fig. 3.8 (2nd row) - albeit less significantly, because the vertex resolution degrades for very background like events. A more plausible explanations might be insufficiencies in the modeling of the optical properties of the bulk ice in the regions of the secondary dust layers (c.f. Sec. 2.2). The level of imperfection is sufficiently small to be disregarded for the purpose of this work. It will therefore be sufficient to only consider the impact of global (systematic) variations of bulk ice optical

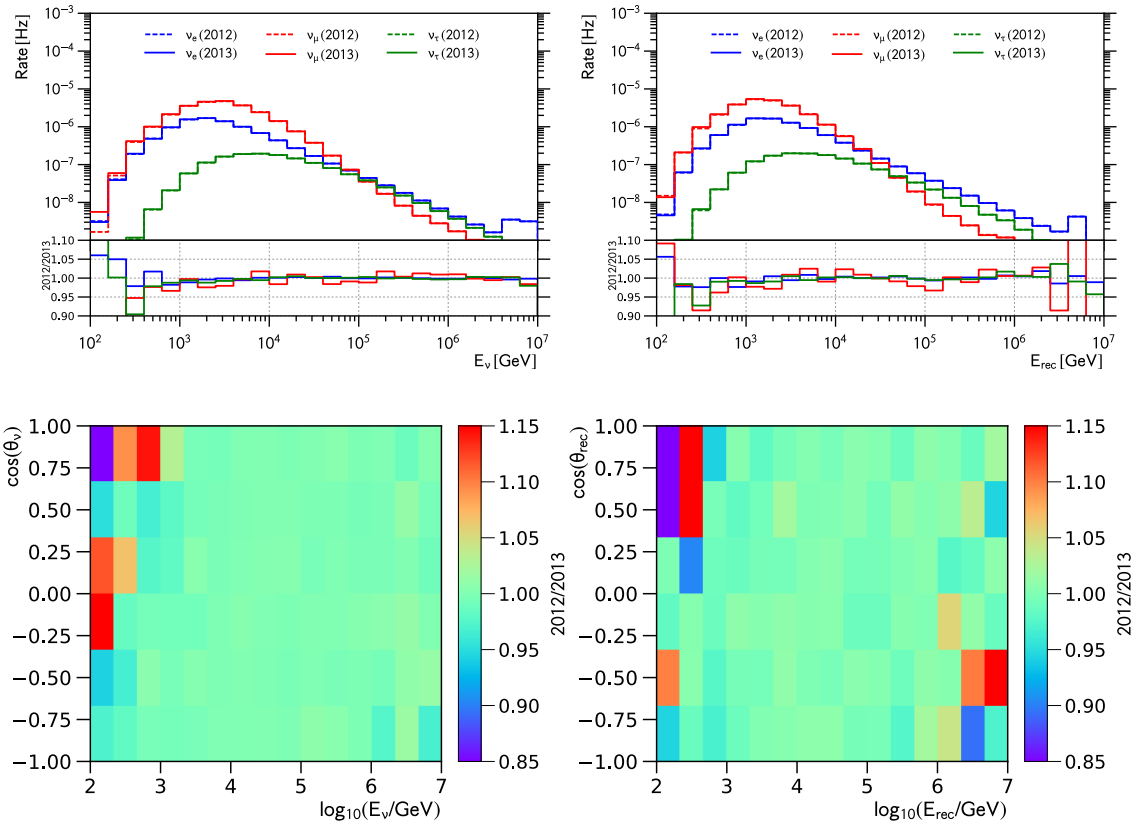


FIGURE 3.26: Comparison between energy (and zenith) distributions predicted from NuGen MC for the years 2012 and 2013. True neutrino energy (top left), reconstructed energy (top right), 2D true neutrino energy and true zenith (bottom left) and 2D reconstructed energy and zenith (bottom right). Essentially no differences are observed.

properties (c.f. Sec. 5.7) during the final analysis. Current computational limitations concerning simulation production render it impossible to properly quantify the influence of local variations of the bulk ice on this analysis.
Since we are interested in measuring the astrophysical neutrino-energy spectrum, we demonstrate in Fig. 3.28 that our simulations accurately model the correlation between total charge collected by all PMTs (Q_{tot}). Fig. 3.28 compares the distribution of the total charge for different thresholds in reconstructed energy: $E_{rec} > 1\,\text{TeV}$ (top left), $E_{rec} > 5\,\text{TeV}$ (top right), $E_{rec} > 10\,\text{TeV}$ (bottom left) and $E_{rec} > 30\,\text{TeV}$ (bottom right). Finally, the selection efficiency with respect to the cascade-L3 filter for the different neutrino flavors as function of neutrino energy is given in Fig. 3.28 (bottom). The selection efficiency peaks at around $E_\nu \sim 100\,\text{TeV}$, where it approaches $\sim 75\%$ for electron and tau neutrinos (slightly larger for the former). While similar efficiency is obtained for muon neutrinos with neutral-current interactions (cascade-like), the efficiency for muon neutrinos with charged-current interactions (starting tracks) is suppressed to $\sim 25\%$, as desired.

3.2.7 Level 5C: The Starting Track Control Sample

The starting track control sample is defined by eq. (3.30) and contains ~ 1600 events that are vastly dominated by charged-current muon neutrino interactions (purity of $\sim 95\%$). This is shown in Tab. 3.11 which summarizes the observed (data) and predicted (simulation) passing rates for this sample. Fig. 3.29 shows the relevant observable distributions: reconstructed energy (top left), total charge (top right), starting track charge (2nd row, left), reconstructed zenith angle (2nd row, right), reconstructed vertex z-position (bottom left) and reconstructed vertex XYscale (bottom right). As expected for starting tracks, we find a larger (relative) contribution from events with large Starting Track Charge (photons from out-going muon) than observed before application of the GBDT classification (Fig. 3.10, 2nd row). The selection efficiencies with respect to the cascade-L3 filter for the different neutrino flavors (and interaction types for ν_μ) are compared in Fig. 3.30 (top). The efficiency for cascade-like events (ν_e, ν_τ, ν_μ-NC) is essentially $\sim 0\%$, meaning that we do not mis-identify cascades as starting tracks,

Rate in Year	**2015** [μHz]	**2014** [μHz]	**2013** [μHz]	**2012** [μHz]
data	$\mathbf{39.7 \pm 1.1}$	$\mathbf{38.4 \pm 1.1}$	$\mathbf{40.3 \pm 1.2}$	$\mathbf{38.6 \pm 1.2}$
ν_e (conv.)	c.f. 2013	c.f. 2013	7.16 ± 0.02	7.12 ± 0.02
ν_μ (conv.)	c.f. 2013	c.f. 2013	25.2 ± 0.2	24.8 ± 0.2
ν_e (prompt)	c.f. 2013	c.f. 2013	0.3109 ± 0.0004	0.3102 ± 0.0004
ν_μ (prompt)	c.f. 2013	c.f. 2013	0.0560 ± 0.0003	0.0556 ± 0.0003
ν_e (astro)	c.f. 2013	c.f. 2013	2.027 ± 0.002	2.023 ± 0.002
ν_μ (astro)	c.f. 2013	c.f. 2013	0.419 ± 0.002	0.417 ± 0.002
ν_τ (astro)	c.f. 2013	c.f. 2013	1.403 ± 0.003	1.400 ± 0.003
$\sum \nu_\mu$ (NC)	c.f. 2013	c.f. 2013	13.8 ± 0.2	13.6 ± 0.2
$\sum \nu_\mu$ (CC)	c.f. 2013	c.f. 2013	11.9 ± 0.2	11.7 ± 0.2
atm. μ	c.f. 2013	c.f. 2013	3.0 ± 0.4	2.9 ± 0.4

TABLE 3.10: Passing rates of Level-5B selection criteria: data and simulation (MuonGun, NuGen)

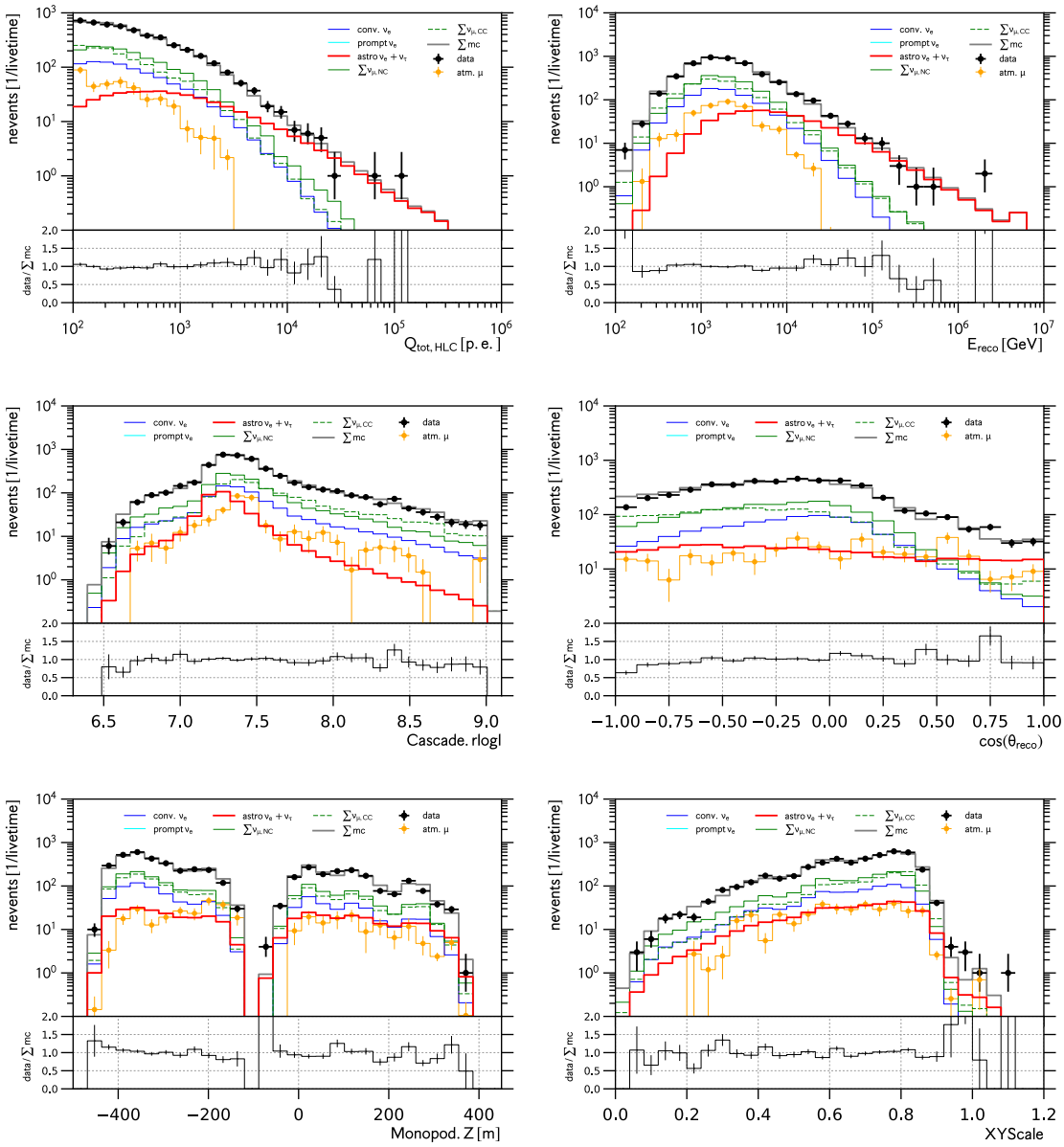


FIGURE 3.27: Observable distributions for L5B selection (cascade signal sample): total PMT charge (top left), reconstructed energy (top right), cascade.rlogl (2nd row, left), reconstructed zenith angle (2nd row, right), reconstructed z-position (bottom left), reconstructed XYscale (bottom right).

as desired. Starting Tracks are best retained (and identified) if they stem from muon neutrinos with (true) energy of $\sim 30\,\text{TeV}$. The peak efficiency for ν_μ-CC events is $\sim 30\%$.

Finally, Fig. 3.31 shows the distribution of the neutrino event type discriminator (starting track score) for the combined neutrino sample (cascade sample and starting track sample). It remains well described by the neutrino simulations at all energies, after muon events have been rejected: all energies (top), $E_{rec} > 1\,\text{TeV}$ (2nd row, left), $E_{rec} > 5\,\text{TeV}$ (2nd row, right), $E_{rec} > 10\,\text{TeV}$ (bottom, left) and $E_{rec} > 25\,\text{TeV}$ (bottom right).

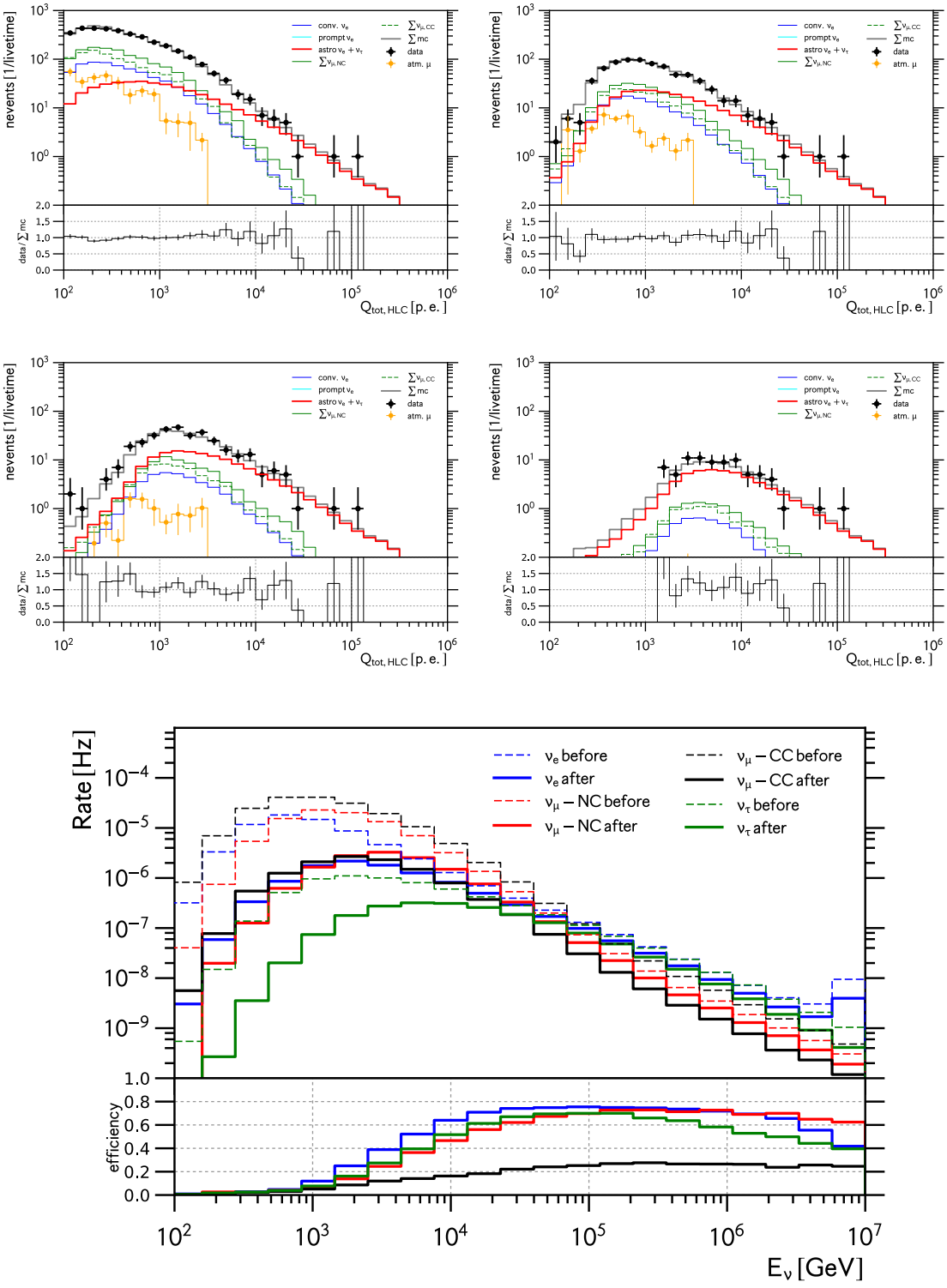


FIGURE 3.28: Distribution of total PMT charge for different energy thresholds. $E_{reco} > 1\,\text{TeV}$ (top left), $E_{reco} > 5\,\text{TeV}$ (top right), $E_{reco} > 10\,\text{TeV}$ (center left) and $E_{reco} > 30\,\text{TeV}$ (center right). Neutrino selection efficiency of L5B cuts w.r.t Cascade-L3 filter (SC) (bottom).

Rate in Year	**2015** [μHz]	**2014** [μHz]	**2013** [μHz]	**2012** [μHz]
data	$\mathbf{13.0 \pm 0.6}$	$\mathbf{13.7 \pm 0.7}$	$\mathbf{13.4 \pm 0.7}$	$\mathbf{12.4 \pm 0.7}$
ν_e (conv.)	c.f. 2013	c.f. 2013	0.071 ± 0.002	0.074 ± 0.002
ν_μ (conv.)	c.f. 2013	c.f. 2013	11.6 ± 0.2	11.8 ± 0.2
ν_e (prompt)	c.f. 2013	c.f. 2013	0.00219 ± 0.00003	0.00230 ± 0.00003
ν_μ (prompt)	c.f. 2013	c.f. 2013	0.0285 ± 0.0002	0.0288 ± 0.0002
ν_e (astro)	c.f. 2013	c.f. 2013	0.0129 ± 0.0002	0.0134 ± 0.0002
ν_μ (astro)	c.f. 2013	c.f. 2013	0.163 ± 0.001	0.165 ± 0.001
ν_τ (astro)	c.f. 2013	c.f. 2013	0.0389 ± 0.0006	0.0394 ± 0.0006
$\sum \nu_\mu$ (NC)	c.f. 2013	c.f. 2013	0.14 ± 0.02	0.15 ± 0.02
$\sum \nu_\mu$ (CC)	c.f. 2013	c.f. 2013	11.7 ± 0.2	11.8 ± 0.2
atm. μ	c.f. 2013	c.f. 2013	0.8 ± 0.2	0.4 ± 0.1

TABLE 3.11: Passing rates of Level-5C selection criteria: data and simulation (MuonGun, NuGen)

3.2.8 Level 5D: The Single Muon Control Sample

In order to constrain the remaining atmospheric muon contribution to the neutrino samples (Sec. 3.2.6 and Sec. 3.2.7) it is necessary to retain a control sample of atmospheric muons with almost cascade-like characteristics, i.e. muons that appear like cascade signal events. The criterion is defined by eq. (3.33) and is satisfied by $\sim$ 10800 events. We estimate that $\sim$ 70% of the events retained in this control sample are muon background, of which $\sim$ 90% enter the detector as single muons. The reconstructed energy spectrum of such events is shown in Fig. 3.32 (top). Some discrepancies are observed at lowest energies and are interpreted to be due to the muon energy threshold of 500 GeV that is used for the single muon simulation (c.f. Sec. 2.5). Otherwise simulation and data agree reasonably well. Fig. 3.32 also shows the distribution of the total charge (2nd row, left). We observe an excess of events for $Q_{tot} > 1000$ p.e., indicative of residual contamination from muon bundles. Since the relative contribution of such events to this control sample is negligible, this can be ignored. In general the agreement between the data and simulations, and thus the quality of our modeling, appears not quite as good as that of the neutrino samples (c.f. Sec. 3.2.6 and 3.2.7). This is partially because of the difficulties involved in generating sufficient muon background simulation, that can lead to large Monte-Carlo statistical uncertainties in parts of the parameter space, the residual muon bundle contamination and finally, possibly more important, due to systematic uncertainties in the modeling of the detector. Since atmospheric muons are strongly suppressed in our neutrino samples, small imperfections in the modeling of muon background will not be resolvable.

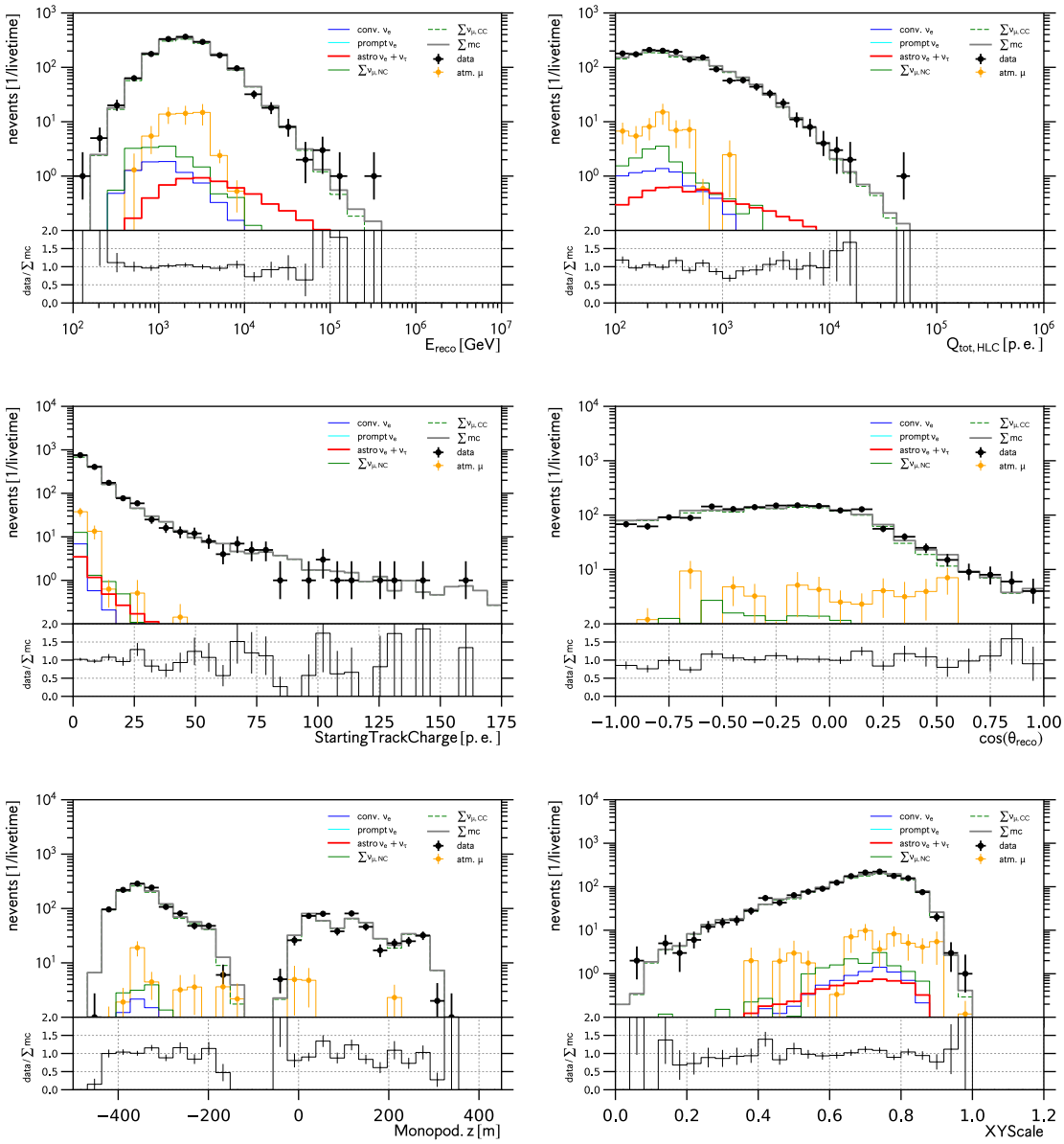


FIGURE 3.29: Observable distributions for L5C selection (starting track control sample): reconstructed energy (top left), total PMT charge (top right), starting track charge (2nd row, left), reconstructed zenith angle (2nd row, right), reconstructed z-position (bottom left), reconstructed XYscale (bottom right).

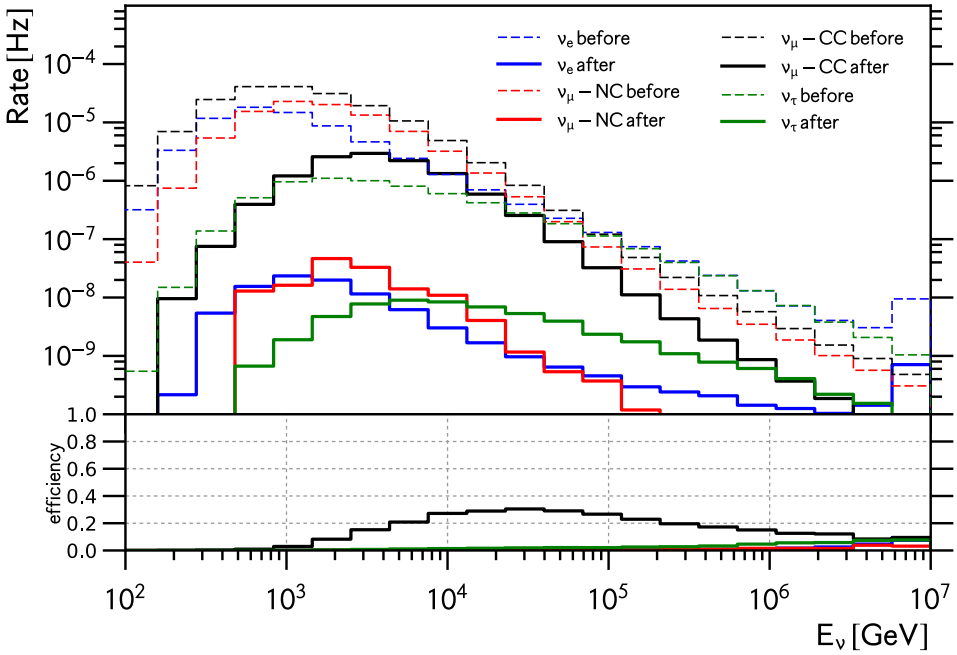


FIGURE 3.30: Neutrino selection efficiency for L5C selection (starting track control sample) w.r.t Cascade-L3 filter (SC).

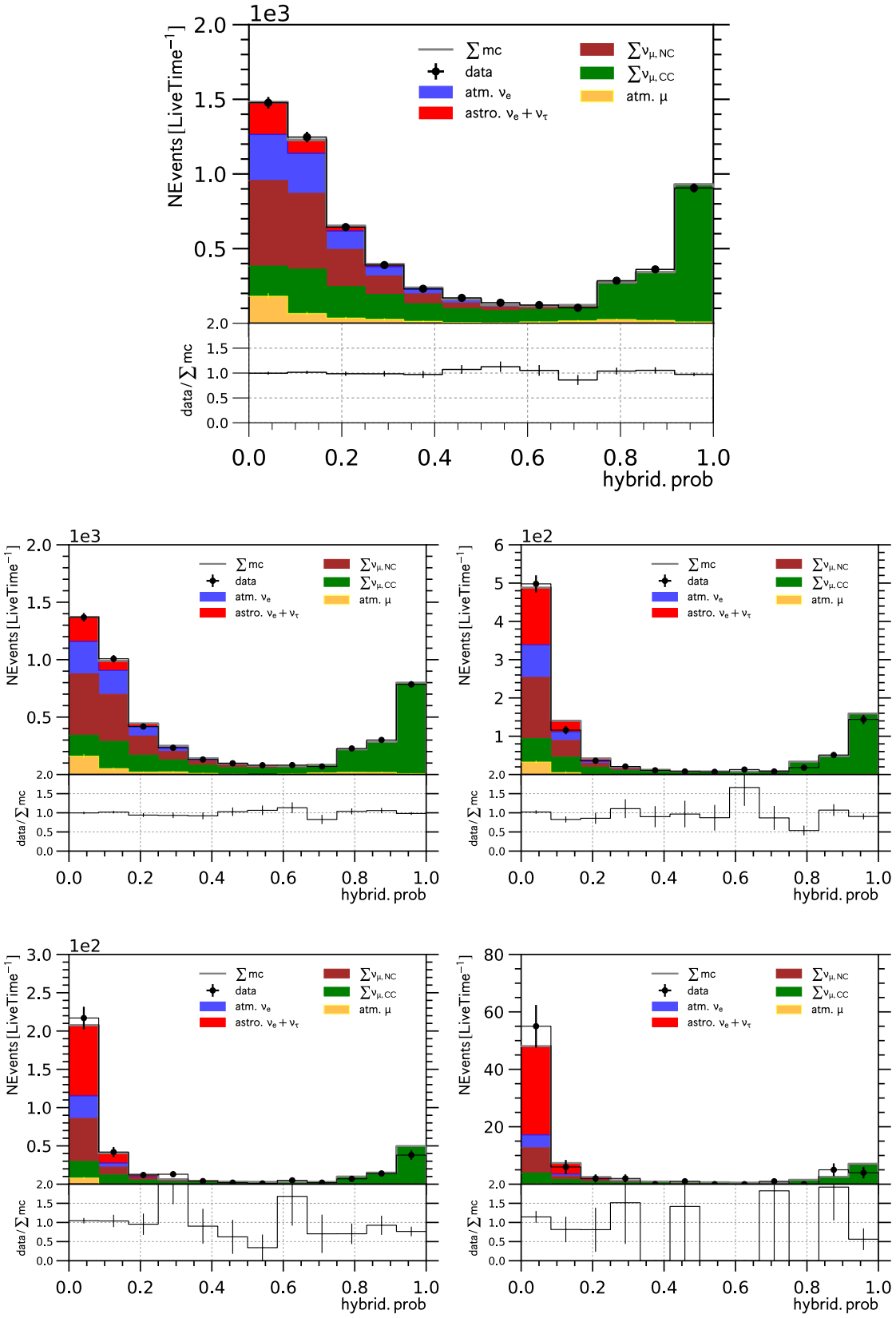


FIGURE 3.31: Distribution of starting track classification score for combined neutrino sample (L5B and L5C). All energies (top), $E_{rec} > 1\,\mathrm{TeV}$ (2nd row, left), $E_{rec} > 5\,\mathrm{TeV}$ (2nd row, right), $E_{rec} > 10\,\mathrm{TeV}$ (bottom left) and $E_{rec} > 25\,\mathrm{TeV}$ (bottom right)

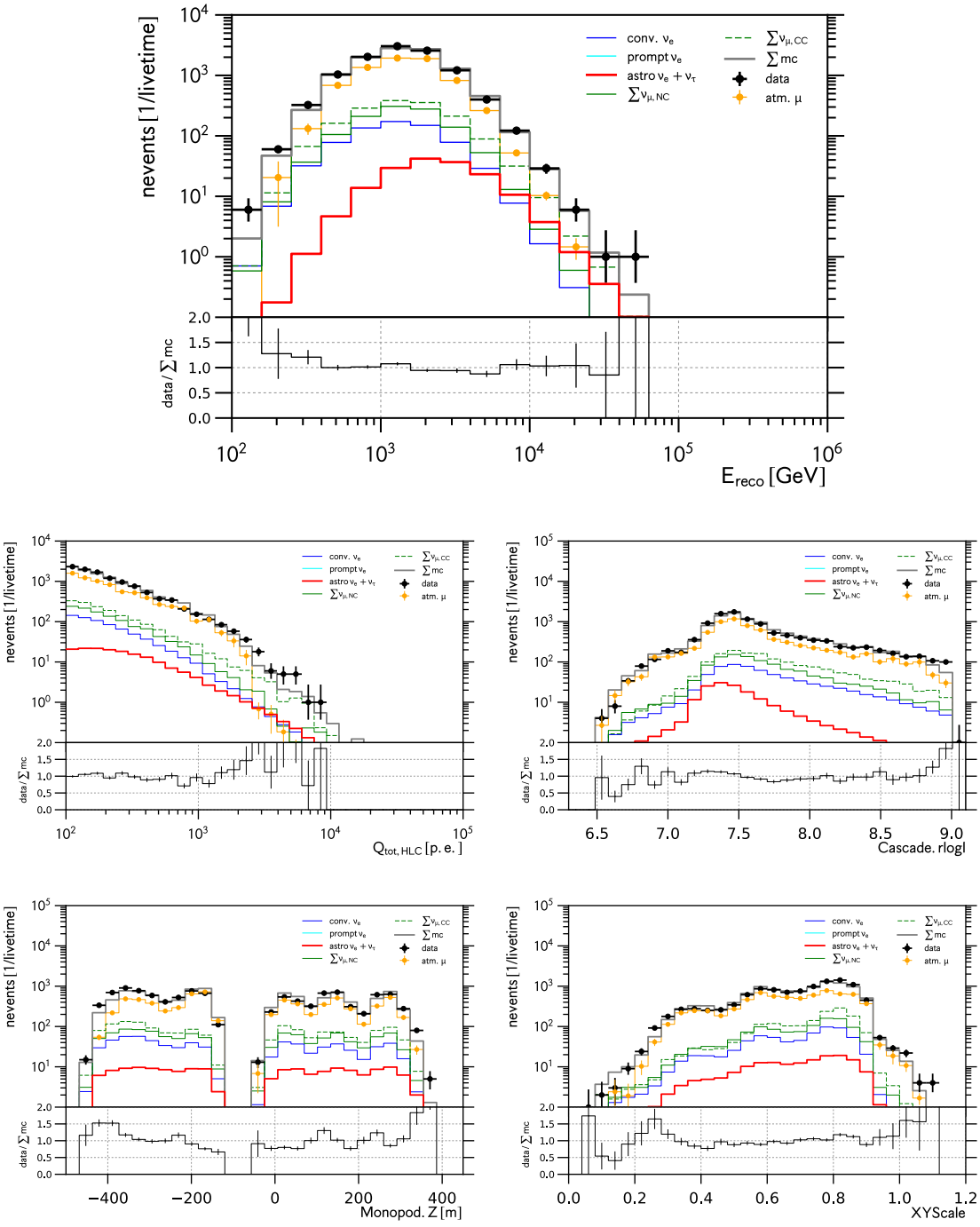


FIGURE 3.32: Observable distributions for L5D selection (single muon control sample): reconstructed energy (top), total PMT charge (2nd row, left), Cascade.rlogl (2nd row, right), reconstructed z-position (bottom left), reconstructed XYscale (bottom right).

3.2.9 Comparison with Previous IceCube Analyses

Since the original discovery of the high energy astrophysical neutrino flux using events above reconstructed energies of $\sim$ 60 TeV [5], two IceCube analyses have studied the spectrum with a lower energy threshold: Cascades [155] above $E_{rec} = 10$ TeV and Medium Energy Starting Events (MESE) [164] above $\sim$ 1 TeV. The event selection, developed in this work, achieves a neutrino purity comparable to the two previous analyses, but increases the detection efficiency for neutrinos with cascade-like signature, especially with energies below 100 TeV. This can be quantified using the effective area for electron neutrinos. This is shown in Fig. 3.33 for all three selections: this work (black), Cascades (10/11) [155] (green) and MESE (red) [164]. Compared to [164], this work increased the electron neutrino effective area by a factor of ~ 3.5 (> 10) for neutrino energies of 10 TeV (1 TeV). At much higher energies of $\sim$ 1 PeV all selections behave essentially identically. At even higher energies (several PeV) this selection performs worse, most notably at the glashow resonance.

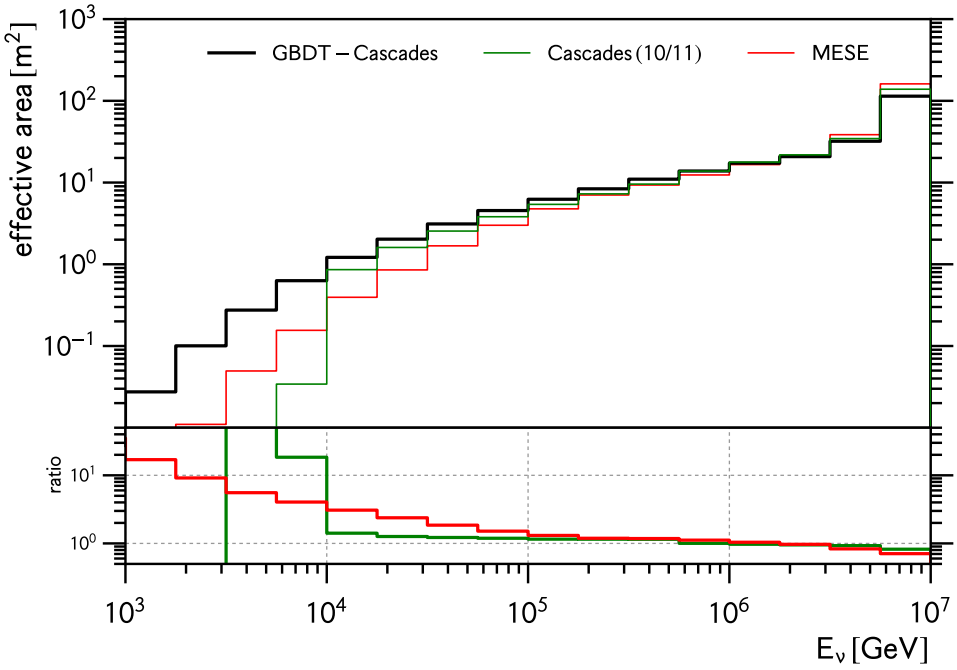


FIGURE 3.33: Comparison of electron neutrino effective areas obtained in this work (black), the previous cascade analysis [155] (green) and Medium Energy Starting Events [164] (red)

3.3 Final Data Sample: Combining with the High Energy Selection

It turns out that the cascade detection efficiency at highest energies can be improved, especially at the top of the detector where this analysis is limited by the pre-cuts (c.f.

Secs. 3.1.3 - 3.2.3). A dedicated event selection, targeting high energy (> 30 TeV) cascades with straight cuts, has been developed by Yiqian Xu at Stony Brook University [153]. As shown in Fig. 3.34 the dedicated high-energy (HE) cascade selection similar cascade signal efficiency (solid blue line) to the GBDT selection (dashed blue line) starting at reconstructed energies of $\sim$ 60 TeV and provides increased efficiency at higher energies. At such energies the Cascade-HE selection has been found to be free from atmospheric muon background [153]. At energies of 1 PeV (10 PeV) the cascade selection efficiency is increased from 60% (40%) to 80% (80%). The suppression of charged-current muon neutrinos (red) is similar to the one obtained in this work. Because of the large overlap (> 90%), expected from simulations, we combined both selections as follows. Events with reconstructed energies below 60 TeV are subjected to the selection criteria discussed in the previous sections. Events above 60 TeV instead must satisfy the criteria defined in [153]. Hence the final starting track and atmospheric muon control samples only contain events with reconstructed energies below 60 TeV. The cascade signal sample, instead covers the entire energy range. As with all other selection criteria, this decision was made before the final sample was revealed (90% of total livetime). The combined selection efficiencies for the different neutrino flavors (and interaction types for ν_μ) as function of neutrino energy are shown in Fig. 3.35 (top) and can be compared to the ones obtained in this work (Fig. 3.28 (bottom)). The combined cascade sample now achieves IceCube's largest effective area for electron neutrinos at all energies. This is shown in Fig. 3.35 (bottom). Compared to previous works we increased the electron neutrino effective area by at least 25%. At lowest energies the increase is larger than a factor of 10. A comparison between the effective areas for the different neutrino flavors is given in Fig. 3.36 (top left). Fig.

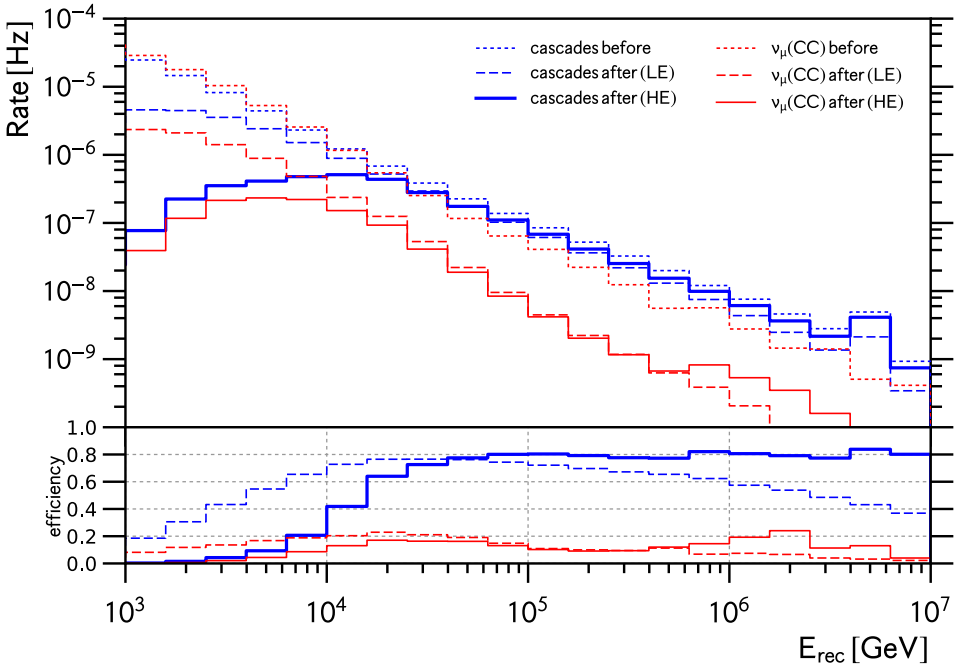


FIGURE 3.34: A comparison of the neutrino selection efficiencies obtained in this work (dashed) to the one obtained with a dedicated high energy cascade selection [153]. The latter retains cascade more efficiently starting at neutrino energies of $\sim$ 100 TeV.

3.36 also shows the declination dependence of the effective area for the different neutrino flavors: ν_e (top right), ν_μ (bottom left) and ν_τ (bottom right). Effective areas for up-going neutrinos are suppressed significantly at high energies $> 100\,\text{TeV}$ due to neutrino absorption in the earth. Tau neutrinos are less effected thanks to tau regeneration in the earth: ν_τ-interactions (CC+NC) produce tau neutrinos in the final states, either directly (NC) or after the decay of the tau-lepton (CC).

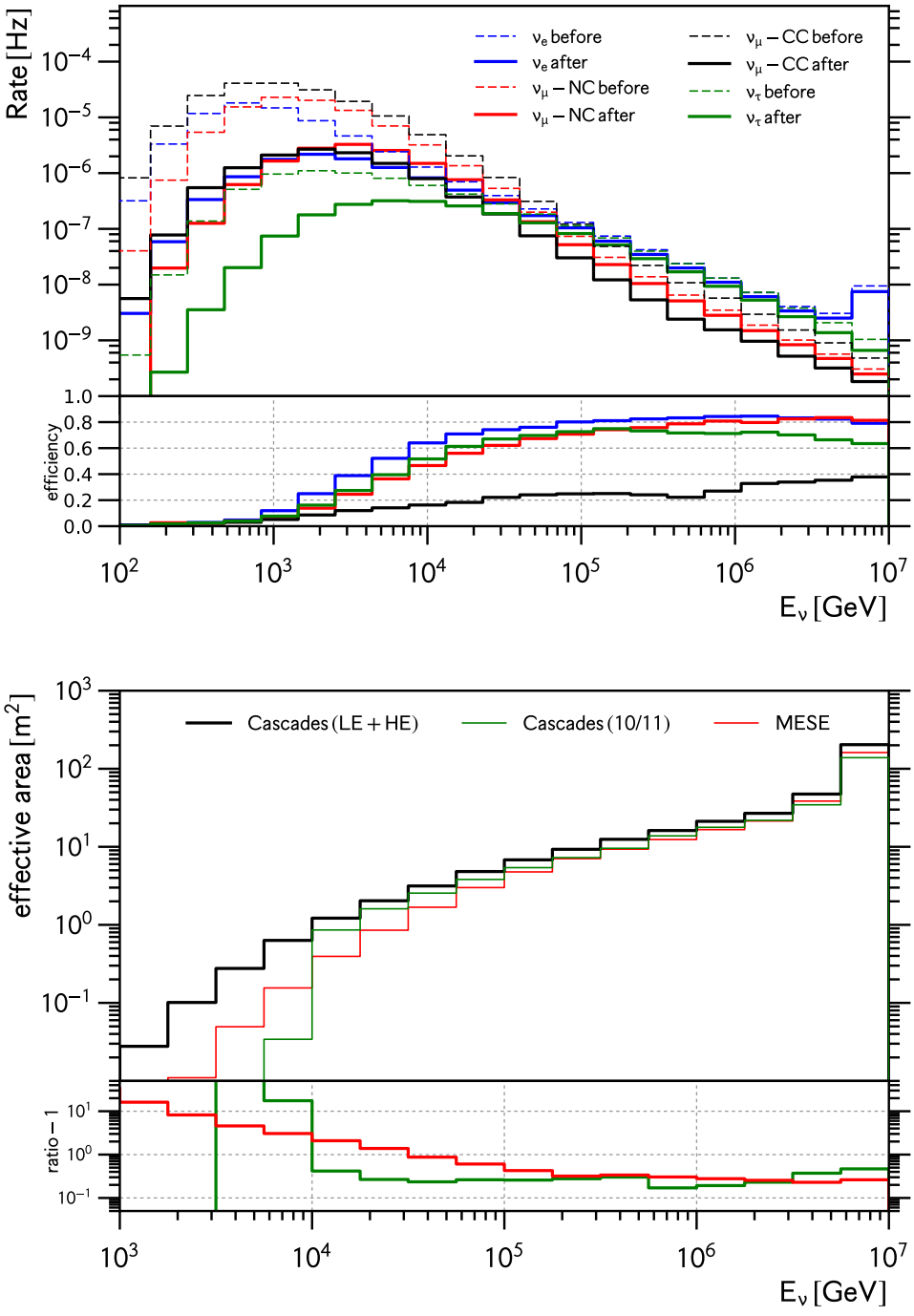


FIGURE 3.35: Combined Cascade Sample (LE + HE selections): neutrino selection efficiency (top) and electron neutrino effective area (bottom): our work (black, previous cascade analysis (green) [155] and MESE (red) [164]. The combined analysis provides largest electron neutrino effective area at all energies.

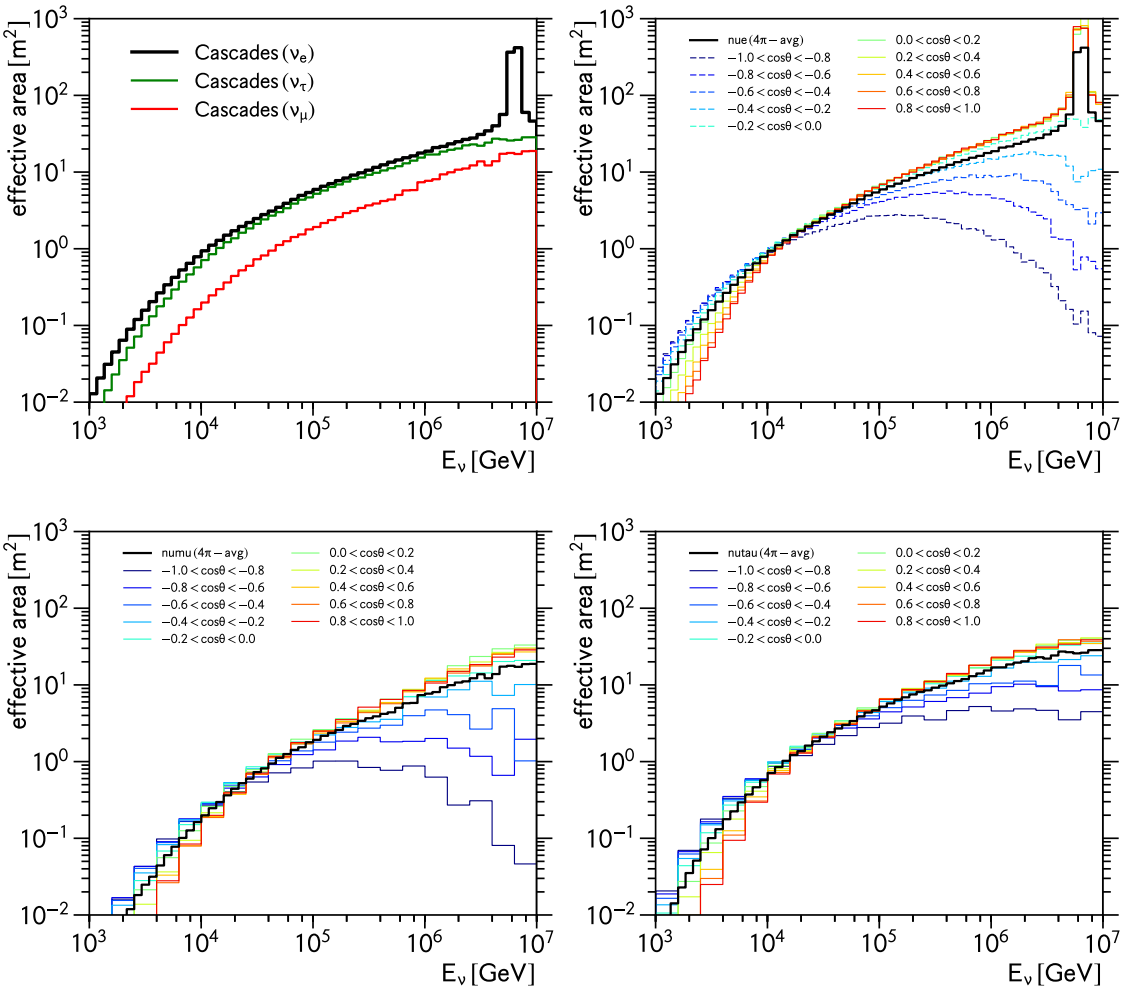


FIGURE 3.36: Combined Cascade Selections (LE+HE): neutrino effective areas for all three neutrino flavors (top left). zenith dependence of electron neutrino effective areas (top right). zenith dependence of muon neutrino effective areas (bottom left). zenith dependence of tau neutrino effective areas (bottom right).

Chapter 4

Systematic Uncertainties

Systematic uncertainties can mainly be classified into those arising from the detector itself, e.g. its response to the Cherenkov light emitted by charged particles traversing it, and those arising from imperfect knowledge of the incident particle flux, i.e. atmospheric muons and neutrinos.

4.1 Modeling the Detector Response

4.1.1 Relative Photon Detection Efficiency

The relative photon detection efficiency of IceCube's optical modules, called "DOM efficiency" hereafter, can be understood as the in-situ fraction of photons incident on a detector optical module that are converted into photo electrons, i.e. actually being detected. Several different uncertainties contribute: the absolute photon detection efficiency of the PMT, a combination of its quantum and photon collection efficiencies, has been calibrated in the laboratory with relative uncertainty of 7.7% [124]. Additional uncertainties include the light attenuation in the glass pressure housing and the gel used for optical coupling of PMT to the glass, as well as local effects in the ice surrounding the DOM (shadowing of photons by the cable, effects from the refrozen ice itself). In-situ calibration of the photon detection efficiency using minimum ionizing muons result in a total relative uncertainty of 10% [165][137].

The DOM efficiency has two main impacts on this analysis. The first is a threshold effect: events that produce the equivalent of less than 100 p.e. in the detector are discarded. Thus the expected number of events above the charge threshold directly scales with the dom efficiency. Secondly, it affects the reconstructed energies of every event in the data sample. If the true dom efficiency was larger (smaller) than the nominal assumption, the reconstructed energies for every event would appear shifted (biased) towards smaller (larger) energies. We determined these effects from simulations assuming a detector with varying dom efficiencies at discrete points $\epsilon \in \{0.81, 0.90, 0.95, 0.99, 1.08\}$. The resulting energy spectra are shown in Fig. 4.1 for true neutrino energies (left) and reconstructed

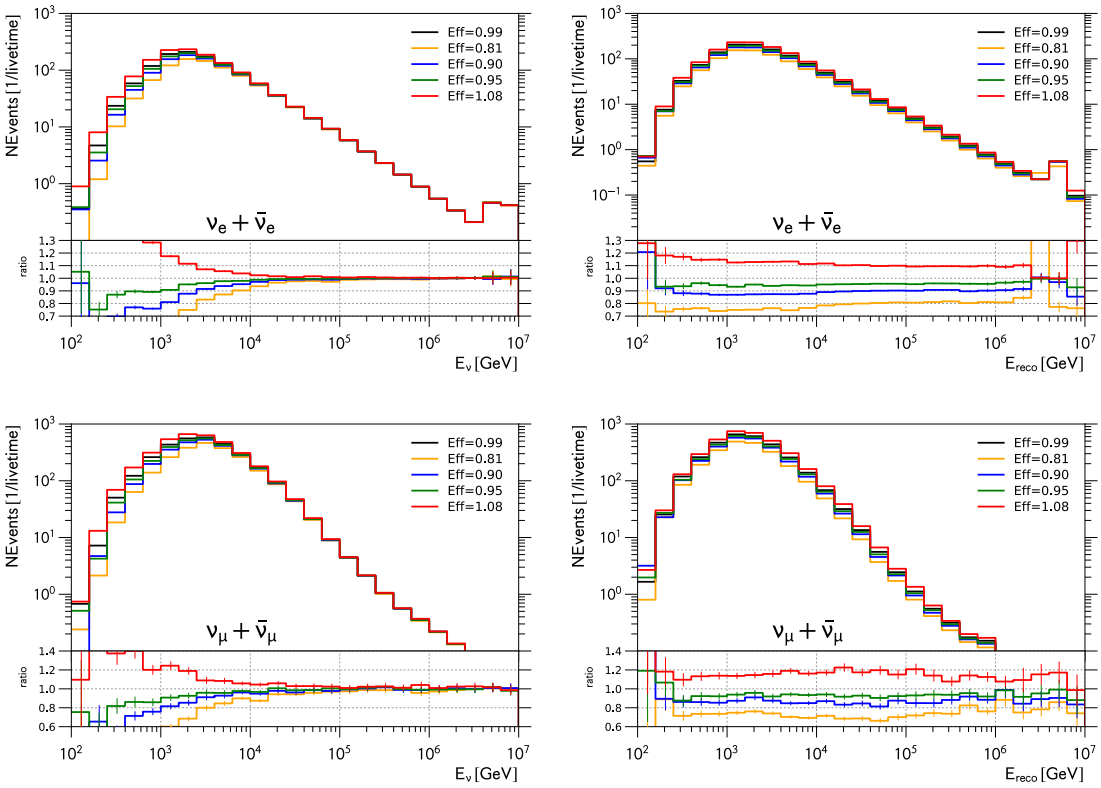


FIGURE 4.1: True (left) and reconstructed (right) energy spectra of electron neutrinos (top) and muon neutrinos (bottom) assuming different true values of the photon detection efficiency of IceCube's optical modules.

energies (right) separately for electron neutrinos (top) and muon neutrinos (bottom). For neutrinos with energies above $\sim 20\,\text{TeV}$ the threshold effect becomes negligible. The impact on reconstructed energies manifests itself mostly as a normalization effect, since the energy spectrum is steeply falling. The strong threshold behavior visible in the true energy spectrum is "smeared out" due to the strong decrease in energy resolution towards lower energies around and below 1 TeV and thus is much less noticeable in the reconstructed energy spectra. There appear to be only minor effects on the true and observed zenith distributions, see Fig. 4.2.

4.1.2 Optical Properties of Glacial Ice at the South Pole

A major contribution to the overall uncertainty in the detector response stems from the imperfect knowledge of the optical properties of the glacial ice that IceCube is embedded in. Variations in the scattering and absorption behavior of the ice can strongly affect the photon transport from the point of emission to the optical modules of the detector. Unfortunately it is computationally prohibitive to study the uncertainty from all the $O(100)$ parameters of the ice-model [125] developed for IceCube. The dimensionality of the problem can be reduced by simplification. Here we investigated the impact of simultaneous changes in scattering and absorption coefficients at all depths relevant for IceCube. The

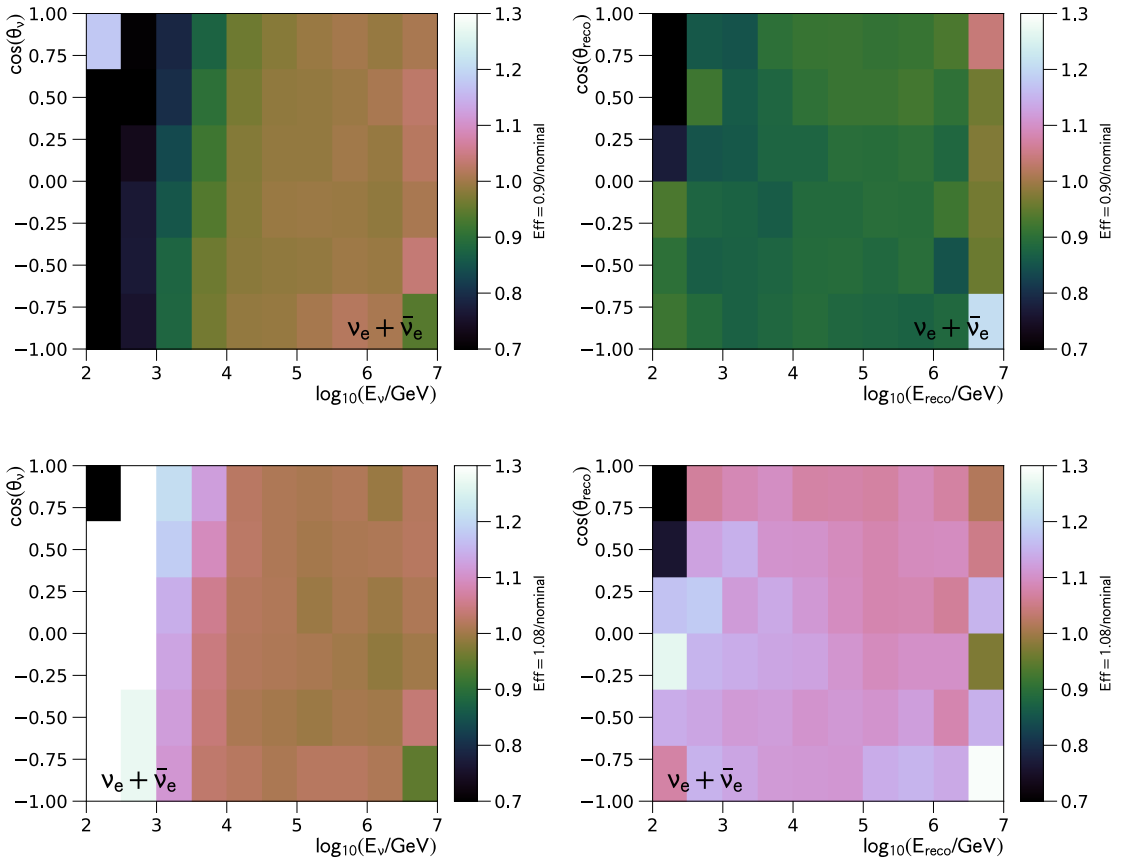


FIGURE 4.2: Ratio between alternative and nominal 2D distributions (energy and zenith) for true (left) and reconstructed (right) quantities for different assumed alternative DOM efficiencies ϵ using electron neutrino simulation. Top: $\epsilon = 0.90/\epsilon = 0.99$. Bottom: $\epsilon = 1.08/\epsilon = 0.99$

uncertainty on global changes of scattering an absorption coefficients has been estimated to be 10% [125].
We simulated three scenarios: an enhancement in scattering across the entire detector by 10%, an enhancement in absorption across the detector by 10% and finally a simultaneous, global reduction in scattering and absorption of 7%. The resulting true (left) and reconstructed (right) energy distributions can be found for electron (top) and muon neutrinos (middle) in Fig. 4.3. Also shown is the effect on the reconstructed zenith distribution (bottom). The phenomenology of systematics related to the ice-properties appears richer than that of the DOM efficiency. We observe non-trivial effects on the energy and zenith distributions. The effect of changing the absorption in the ice is similar to that of reducing the DOM efficiency. Since in this scenario photons are more likely to get absorbed the same event will produce less charge in the IceCube sensors, leading to threshold effects below neutrino energies of $E = 20\,\text{TeV}$ and an under-estimated energy reconstruction. Enhancing the scattering of photons in the ice instead pre-dominantly effects the reconstructed zenith distribution. Compared to the standard scattering coefficients events are more likely to be reconstructed as down-going (i.e. reconstructed to originate from the Southern hemisphere), while the number of events that are reconstructed as up-going (i.e.

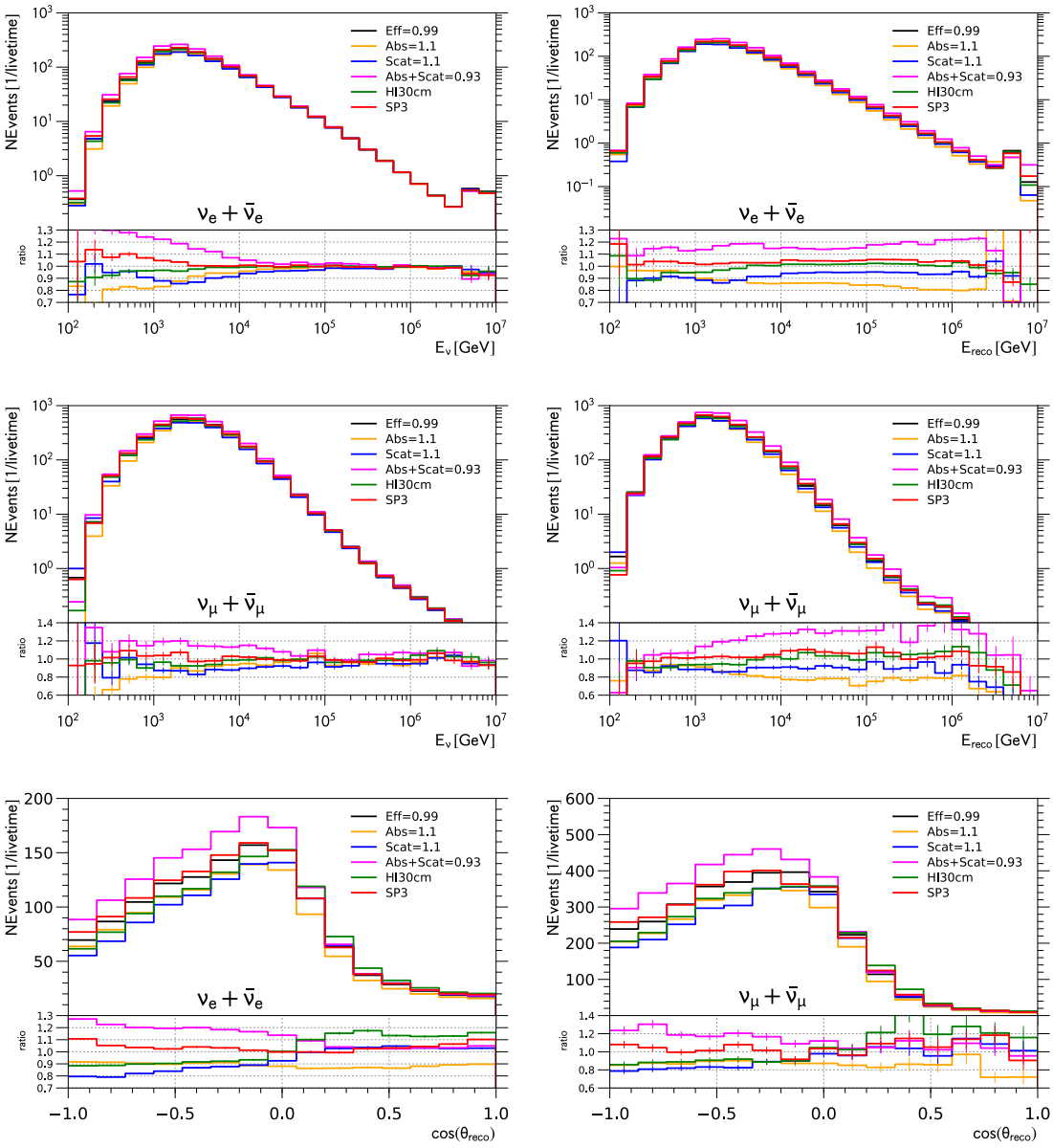


FIGURE 4.3: Global changes in scattering and absorption of photons in the detector by 10%. True (left) and reconstructed (right) energy distributions for electron (top) and muon neutrinos (middle). Zenith distributions (bottom) for electron (left) and muon (right) neutrinos. Also shown is the impact of changes in Hole-Ice parametrization (green).

reconstructed to originate from the Northern hemisphere) is reduced by up to 20%. In addition threshold effects cause a reduction in total number of events of $\sim$ 10% below neutrino energies of 10 TeV.

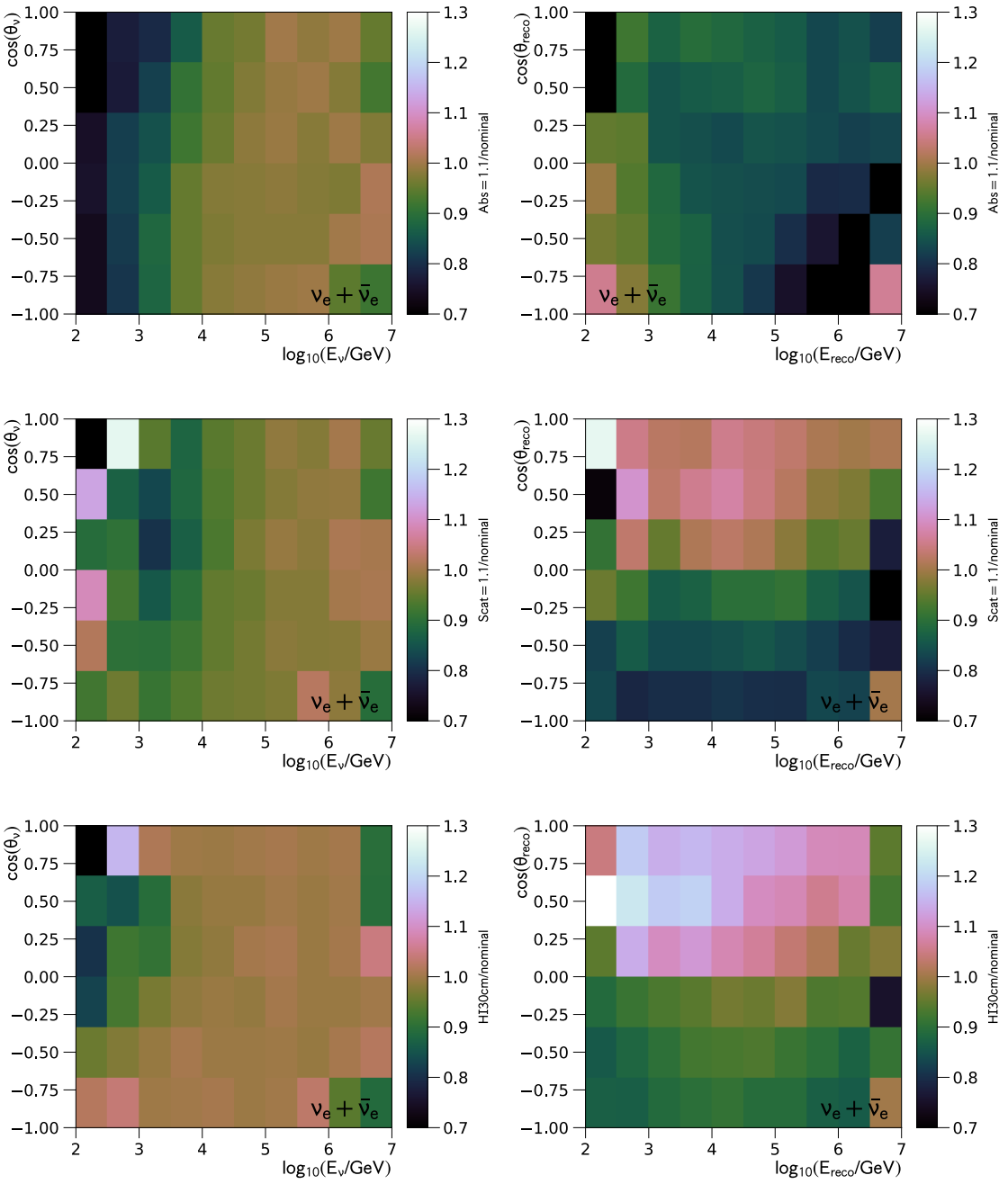


FIGURE 4.4: Global changes in absorption (top) and scattering (center) of photons in the detector by 10%. Ratios between alternative simulation and nominal simulation. True (left) and reconstructed (right) 2D energy and zenith distributions. Bottom: Different angular DOM acceptance corresponding to 30 cm scattering length.

4.1.3 Optical Properties of Hole Ice

The term "Hole Ice" refers to the refrozen ice in the drill holes that were melted to lower IceCube's instrumentation into the ice. Those holes have a diameter of roughly

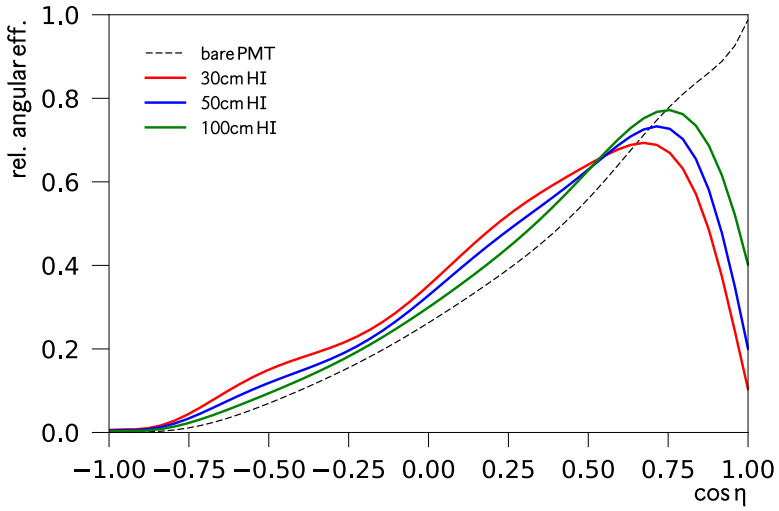


FIGURE 4.5: Relative Photon Efficiency of the IceCube DOM as function of the incident angle of the photon w.r.t to the PMT for different hole-ice scattering lengths: 30 cm (red), 50 cm (blue), 100 cm (green) and no hole-ice scattering (lab measurement) in black.

60 cm. Camera observations suggest the existence of column of air-bubbles confined to a cylinder of $\sim$ 8cm diameter that were formed during the process of refreezing [166]. These bubbles strongly increase the scattering of light in the vicinity of the IceCube DOMs, thus effectively altering the angular acceptance of DOMs in-situ compared to the bare DOM. The photon efficiency for head-on illumination is reduced while the efficiency for other direction is enhanced compared to the angular efficiency calibrated in the laboratory. Nominal IceCube simulations assume an in-situ angular acceptance corresponding to a Hole Ice column of 50 cm scattering length that was measured with IceCube's predecessor experiment AMANDA. Alternative simulations assume scattering lengths of 30 cm and 100 cm. Fig. 4.5 shows the corresponding relative angular efficiencies (solid) of the IceCube DOM compared to the one calibrated in the lab (dashed). The effect of changing the angular acceptance from 50 cm to 30 cm on the final sample observable distributions is shown in Fig. 4.3 (green). The true and reconstructed energy distributions are mostly insensitive to this change except for the lowest energies below a few TeV. The effect on the zenith distribution is similar to an enhancement in bulk ice scattering: the number of reconstructed down-going events is increased by $\sim$ 20% compared to the nominal simulation. The effect remains visible at highest energies above 100 TeV.

4.2 Modeling the Background from Atmospheric Neutrinos

4.2.1 Conventional Atmospheric Neutrinos

Impact of Hadronic Interaction Models

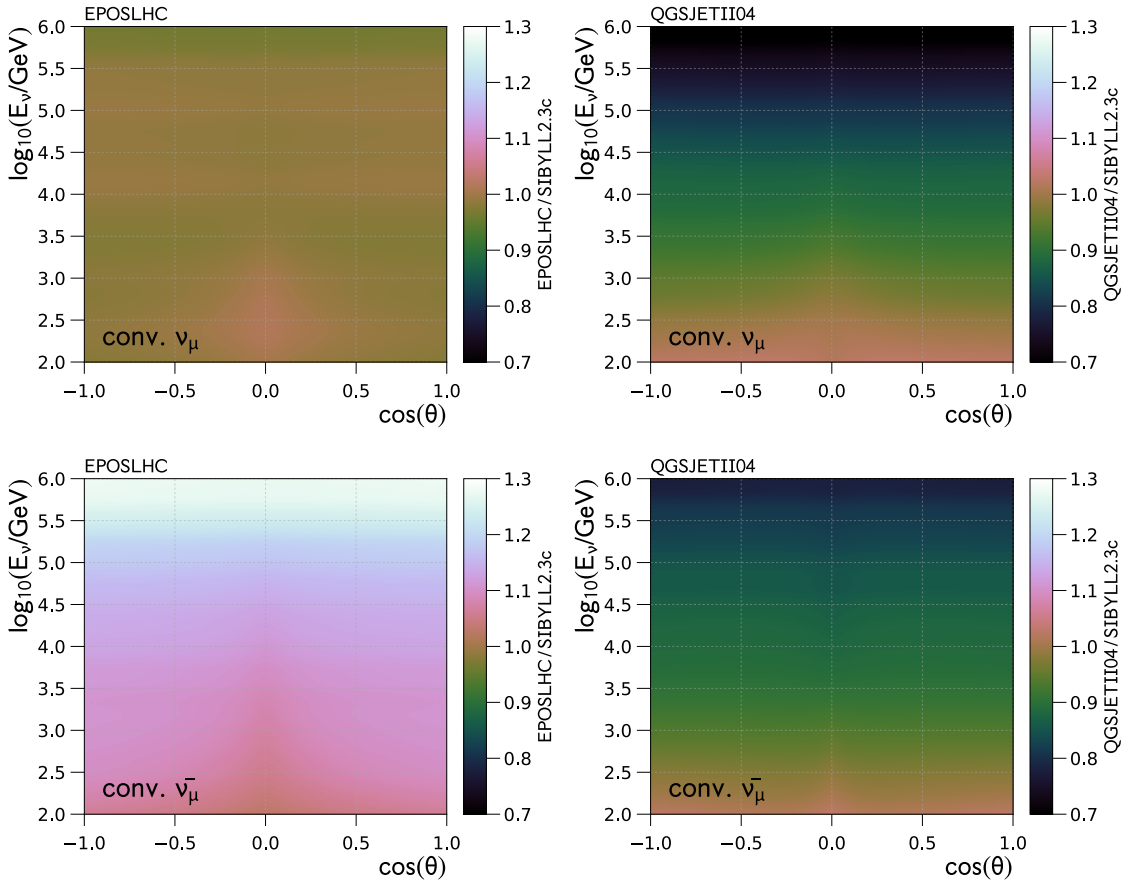


FIGURE 4.6: Ratio between predicted conventional muon neutrino (top) and muon antineutrino (bottom) fluxes for different hadronic interaction models. EPOS-LHC / SIBYLL-2.3c (left), QGSJETII-04 / SIBYLL-2.3c (right)

The impact of uncertainties in the modeling of hadronic interactions during cosmic ray air-shower development on atmospheric neutrino flux predictions has been studied in [167]. Depending on energy differences among the different hadronic interaction models lead to differences in the atmospheric neutrino flux of order $20\% - 30\%$. Since then those interaction models have been refined to take into account experimental data obtained with the Large Hadron Collider leading to better agreement in the context of air-shower modeling [168]. To estimate the corresponding uncertainties on the conventional atmospheric neutrino flux from the decays of pions and kaons we use MCEq [169] and calculate the neutrino flux for fixed primary cosmic-ray flux and atmosphere model but for different hadronic interaction models that include LHC data.

- Primary Cosmic Ray Flux Model: Gaisser-Hillas 2012, "H3a" [170]
- Atmospheric Density Model: Corsika Atmosphere 17 (U.S. Std. Atmosphere) [171][172]
- Hadronic Interaction Models:
  SIBYLL 2.3c[173], QGSJETII-04[174],
  EPOS-LHC[175], DPMJETIII-17.1[176][177]

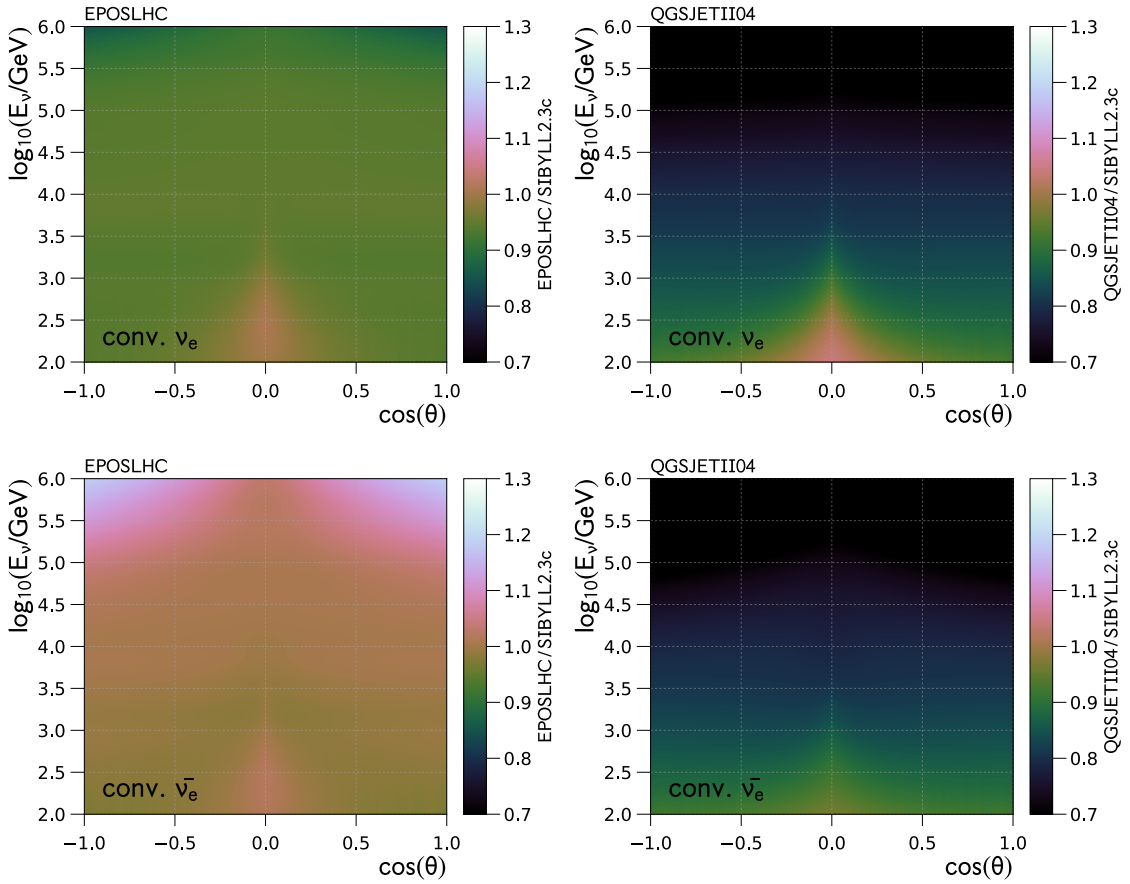


FIGURE 4.7: Ratio between predicted conventional electron neutrino (top) and electron antineutrino (bottom) fluxes for different hadronic interaction models. EPOS-LHC / SIBYLL-2.3c (left), QGSJETII-04 / SIBYLL-2.3c (right)

The left (right) plots in Figs. 4.6 and 4.7 show the ratios between the fluxes predicted by EPOS-LHC (QGSJETII-04) and the SIBYLL 2.3c prediction for atmospheric muon and electron (anti-)neutrinos, respectively. The different modals impact the energy dependence of the flux more than it's zenith dependence.

We studied the effect on the observable distributions in the final data sample of this analysis using Monte Carlo simulations of neutrino propagation, neutrino interaction, taking into account the different interaction cross-sections for neutrinos and anti-neutrinos, and detector response to the emitted Cherenkov light. The difference in the atmospheric flux due to different hadronic interaction modeling is accounted for by adjusting the weights of each individual MC event accordingly. Fig. 4.8 shows the reconstructed energy (top) and zenith (middle, bottom) distributions for the muon neutrinos, $\nu_\mu + \bar{\nu_\mu}$ (left) and electron neutrinos, $\nu_e + \bar{\nu_e}$ (right) for the different hadronic interaction models mentioned above (colors). Also shown, for reference, is the baseline flux used in IceCube, modified HKKMS06[99] (black). All distributions have been corrected for the atmospheric self-veto effect (cf. Sec. 4.2.3).
For muon neutrinos we find variations of $\sim 10\%$ for energies up to 1 TeV and $\sim 20\%$ for higher energies up to several tens of TeV. At highest energies above 100 TeV uncertainties

become larger than 30% but at these energies the conventional atmospheric flux is subdominant compared to neutrinos of astrophysical origin und thus uncertainties become irrelevant for the purpose of this dissertation. The uncertainties in the electron neutrino prediction due to the different interaction models appear larger than for muon neutrinos. Finally it is worth observing that the differences in the muon neutrino expectation between all four MCEq calculations and IceCube's baseline model are independent of the zenith angle and only functions of energy. For electron neutrinos instead we observe an increasing difference towards vertical trajectories. This difference, however, will hardly be observable, since the electron neutrino contribution to atmospheric neutrinos is smaller than that from muon neutrinos and other detector related systematic uncertainties (e.g. scattering in bulk and hole ice) effect the directional reconstruction of both components.

Impact of Primary Cosmic Ray Flux

At neutrino energies relevant for this work the largest contribution to the uncertainties in the calculation of conventional atmospheric neutrino fluxes are hadronic interaction models [167]. Uncertainties related to the primary cosmic ray flux, however, are non-negligible. Here we simplify the problem and only study the impact of global changes in the assumed spectral index of the primary cosmic rays (shape uncertainties). Uncertainties related to the normalization of the flux will be taken into account in the analysis by leaving it as a free parameter in the fit. The effect of varying the spectral index of the primary cosmic-rays on the observed neutrino spectrum is shown in Fig. 4.9 ($-0.05 \leq \Delta\gamma_{CR} \leq +0.05$). Across the relevant energy range, the relative change in the predicted neutrino spectrum is $\pm 20\%$ (at $\sim 100\,\mathrm{TeV}$ compared to $\sim 1\,\mathrm{TeV}$).

4.2.2 Prompt Atmospheric Neutrinos

A selection of recent predictions of the prompt neutrino flux is shown in Fig. 1.11. At energies relevant for this work the predicted fluxes agree rather well, especially in shape. Uncertainties, for example related to the heavy quarks production cross sections, mostly affect the overall normalization, which we will account for by leaving it as a free parameter in the fit. Uncertainties related to the primary cosmic-ray flux are similar to the ones observed for conventional atmospheric neutrinos (c.f. Fig. 4.9). More detailed discussions of relevant uncertainties in the flux calculation can be found in [106][178][105][104].

4.2.3 The Atmospheric Self-Veto Effect

While the flux of atmospheric neutrinos at Earth's surface is approximately symmetric around the horizon in IceCube's local coordinate system, the expected background contribution from atmospheric neutrinos to the final data sample is not. Neutrinos produced in cosmic ray air-showers in the northern terrestrial hemisphere arrive at the detector unaccompanied by atmospheric muons from the same air-showers, because these muons cannot penetrate through the Earth. Neutrinos that originate from the southern hemisphere with sufficiently large energies instead have a non-zero probability of being accompanied by a detectable muon. Event selection criteria designed to isolate pure neutrinos by rejecting

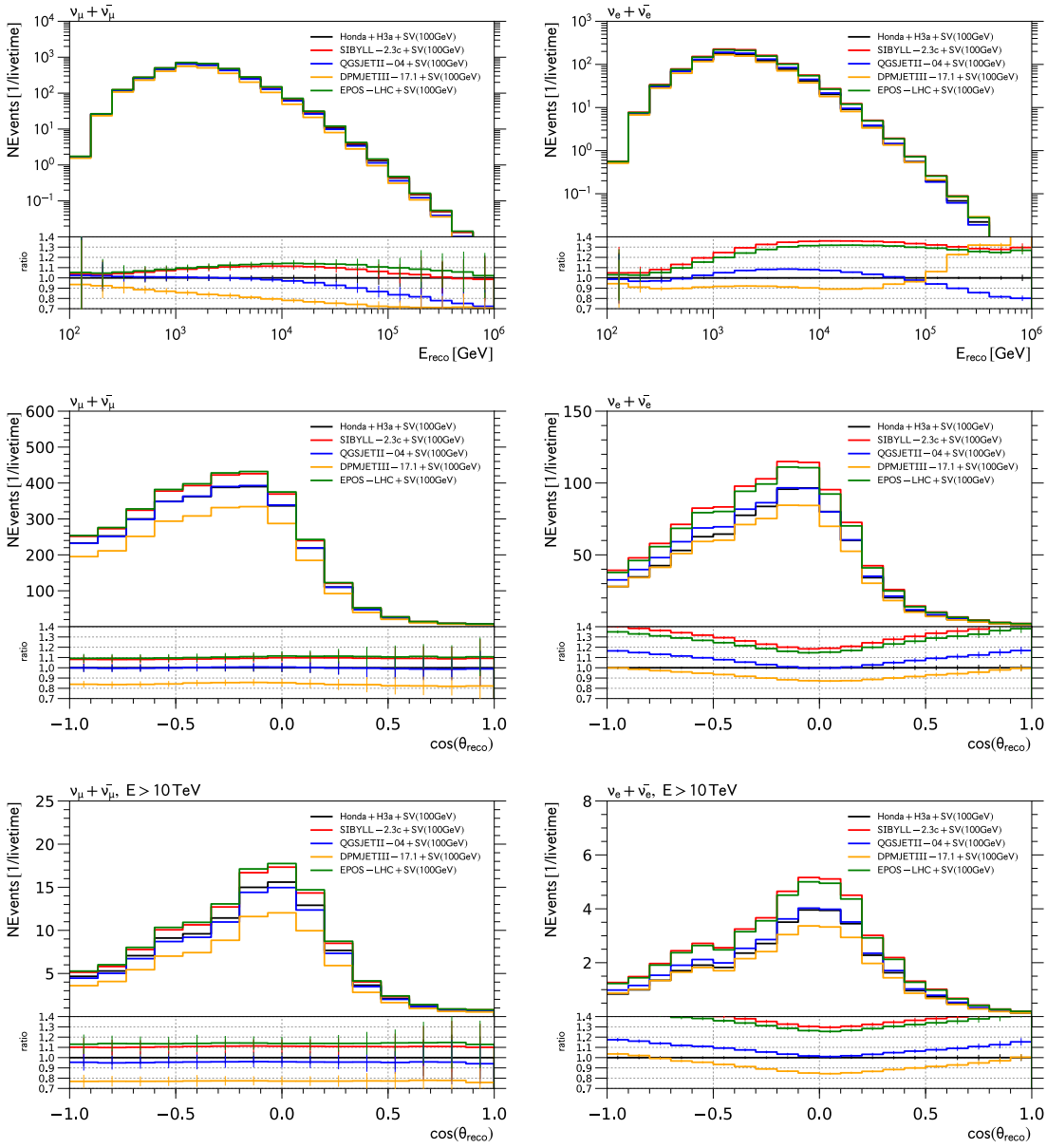


FIGURE 4.8: Reconstructed energy (top) and zenith (middle, bottom) distributions for $\nu_\mu + \bar{\nu_\mu}$ (left) and $\nu_e + \bar{\nu_e}$ (right) assuming different hadronic interaction models.

muon-like events, as those presented in this dissertation, thus suppress the atmospheric neutrino background from the southern sky in comparison to the atmospheric neutrino contribution from the northern sky - even if these criteria themselves are not explicitly dependent on the arrival direction of the event. It follows that astrophysical neutrinos can be measured in the southern hemisphere with a better signal-to-background ratio than in the northern hemisphere, assuming the astrophysical neutrino flux to be isotropic.
This "atmospheric self-veto" effect was first studied in [180] for the case of both particles (muon and neutrino) stemming from the same parent meson by looking at the kinematics of the two-body decays of pions and kaons into muon and muon-neutrinos: $\pi^\pm / K^\pm \rightarrow \mu^\pm + \nu_\mu\,(\bar{\nu_\mu})$. Neutrino and muon can arrive at the detector sufficiently aligned

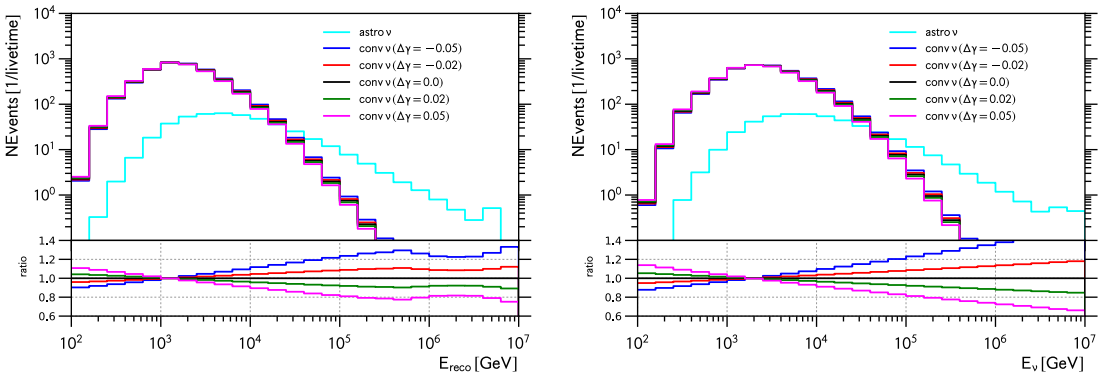


FIGURE 4.9: Reconstructed energy (left) and true neutrino energy (right) distributions for $\nu_\mu + \bar{\nu_\mu}$ (left) and $\nu_e + \bar{\nu_e}$ (right) assuming different values for the spectral index of the primary cosmic-rays

such that they can be considered companions with typical distances between neutrino and muon (after 10 km of propagation) of less than 0.1 m (1 m) for neutrino energies of 10 TeV if the parent meson was a pion (kaon). The probability of obtaining a muon that is able to penetrate through the ice shield into the detector and thus the probability of vetoing the neutrino increases with the neutrino energy but decreases with zenith angle since the ice thickness appears largest for near-horizontal trajectories (see Fig. 4.10 (bottom)). For neutrino energies greater than 10 TeV and zenith angles below 60° ($\cos\theta > 0.5$) veto efficiencies of $\sim 99\%$ were obtained [180].
In addition it is possible that the neutrino arrives at the detector accompanied by a muon that stems from the same air-shower but a different decay. The calculation of [180] has been extended to account for this in [179]. It only slightly increased the veto probability for atmospheric muon neutrinos. However uncorrelated muons from different meson decays provide the only mechanism that suppresses atmospheric electron neutrinos in the southern hemisphere because electron neutrinos are predominantly produced together with electrons. Hence the calculation of [179] can also be applied to atmospheric electron neutrinos.

The parametrization of the veto probabilities obtained in [179] is used within the IceCube Collaboration to predict the contribution from atmospheric neutrino background from the southern hemisphere [5][155][164][181][182] for any given flux of atmospheric neutrinos at the surface. It has been modified following [139] to take into account the depth dependence of the ice overburden (see fig. 4.10 (bottom)). In this model there is only one free parameter $E^{det}_{\mu,min}$, i.e. the minimum energy of a muon at depth that would be identified and removed from the data sample by the event selection. Fig. 4.10 shows for fixed veto threshold $E^{det}_{\mu,min} = 100\,\text{GeV}$ the predicted veto probabilities as function of neutrino energy and zenith angle for electron neutrinos (left) and muon neutrinos (right), originating from π/K decays (top) and heavier mesons involving charm quarks (middle). As discussed above, the veto probability increases with the neutrino energy but decreases with zenith angle and muon neutrinos are more efficiently vetoed than electron neutrinos.

Verifying the Model

The validity of the self-veto parametrization, specifically the choice of $E^{det}_{\mu,min} = 100\,\text{GeV}$,

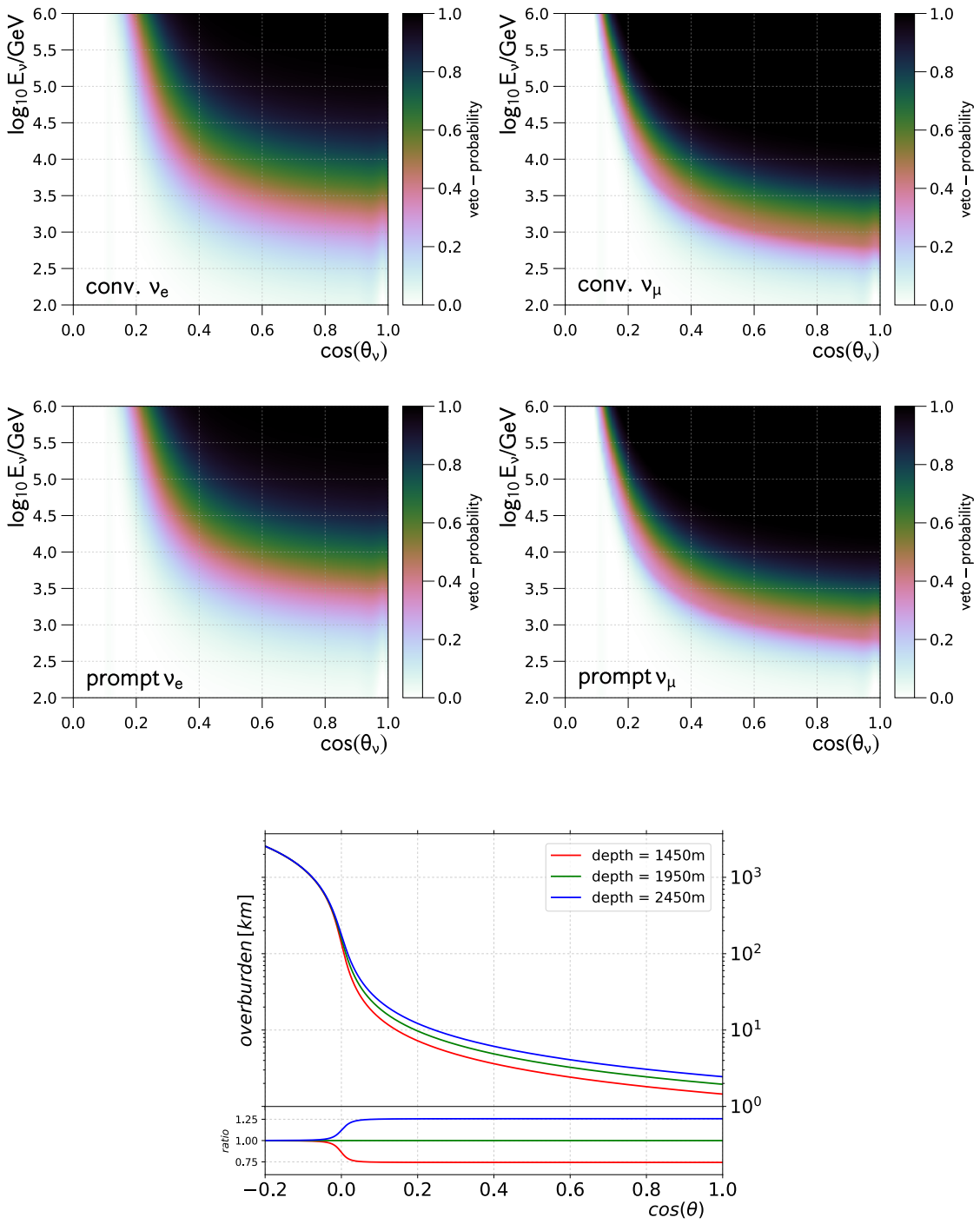


FIGURE 4.10: Veto probabilities for atmospheric neutrinos from the southern hemisphere due to accompanying atmospheric muons according to [179][139] for $E^{det}_{\mu,min} = 100\,\text{GeV}$: conventional ν_e (top left), conventional ν_μ (top right), prompt ν_e (middle left), prompt ν_μ (middle right). Thickness of the ice-shield on top the detector as function of zenith angle (bottom).

can be verified using Monte-Carlo simulations. For this purpose we use full air-shower simulation including neutrino production using CORSIKA [151] and subsequent simulation of particle (neutrino and muon) propagation to the detector, particle interactions as well as the detector response to the corresponding light signals [1]. The simulated events were then filtered by applying all selection criteria starting from the standard IceCube filters to the additional selection criteria developed for this dissertation. This procedure results in the most realistic estimate of the self-veto induced suppression of atmospheric neutrinos as it directly mimics real experimental data.
We will now compare the prediction from the parametrization to the result of the full simulation. For this purpose the parametrized veto probabilities (Fig. 4.10) have to be applied to standard neutrino-only simulations. In order to study the self-veto effect in isolation from other systematic uncertainties, the neutrino-only simulations need to describe the same physics assumptions as those used for the full CORSIKA [151] simulations. More specifically, the assumed primary cosmic-ray flux parametrization of the atmosphere and the description of hadronic interactions within the air-shower need to be identical. The baseline model for the flux of atmospheric neutrinos used in IceCube (HKKMS06 [99]) does not satisfy this requirement. We therefore rely again on MCEq [169] to solve numerically the system of coupled differential equations ("Cascade equations") that govern particle production, interactions, transport and decay in earth's atmosphere to predict the resulting neutrino flux at earth's surface. The model choices that resemble the aforementioned assumptions in the full CORSIKA simulation are given below.

- Primary Cosmic Ray Flux Model: Gaisser-Hillas 2012, "H3a" [170]
- Atmospheric Density Model: Corsika Atmosphere 26 (South Pole atmosphere for July 01, 1997 (MSIS-90-E)) [171]
- Hadronic Interaction Model: SIBYLL 2.3 [173]

The resulting atmospheric neutrino flux $\Phi_j^k\left(E_\nu, \theta_\nu\right)$, is then used to calculate flux weights for each event from the neutrino-only simulation and corrected for the self-veto rejection efficiency $\xi_j^k\left(E_\nu, \theta_\nu\right)$ (Fig. 4.10) where k indexes the production channel (pion/Kaon or charm):

$$\begin{aligned} w_{ij}\left(E_{\nu,i}, \theta_{\nu,i}\right) = & \, w_{ij}^{\pi/K}\left(E_{\nu,i}, \theta_{\nu,i}\right) \times \xi_j^{\pi/K}\left(E_{\nu,i}, \theta_{\nu,i}\right) \\ & + w_{ij}^{\text{charm}}\left(E_{\nu,i}, \theta_{\nu,i}\right) \times \xi_j^{\text{charm}}\left(E_{\nu,i}, \theta_{\nu,i}\right) \end{aligned} \tag{4.1}$$

with i, j indexing the i-th simulated neutrino of type $j \in \{\nu_e, \nu_\mu, \bar{\nu}_e, \bar{\nu}_\mu\}$ with $w_j^k = \Phi_j^k\left(E_\nu, \theta_\nu\right) \times w_j^{\text{gen}}\left(E_\nu, \theta_\nu\right)$ where $w_j^{\text{gen}}\left(E_\nu, \theta_\nu\right)$ are importance weights unrelated to the neutrino flux.

Fig. 4.11 shows the predicted atmospheric (conventional + prompt) neutrino energy spectra after application of all selection criteria obtained from full CORSIKA simulations (red) and neutrino-only simulations with self-veto efficiency corrections corresponding to different choices of $E_{\mu,min}^{det}$ (blue: 100 GeV, black: 400 GeV, green: 700 GeV) for

[1] the simulations were performed by K. Jero and N. Wandkowsky, IceCube, University of Wisconsin-Madison

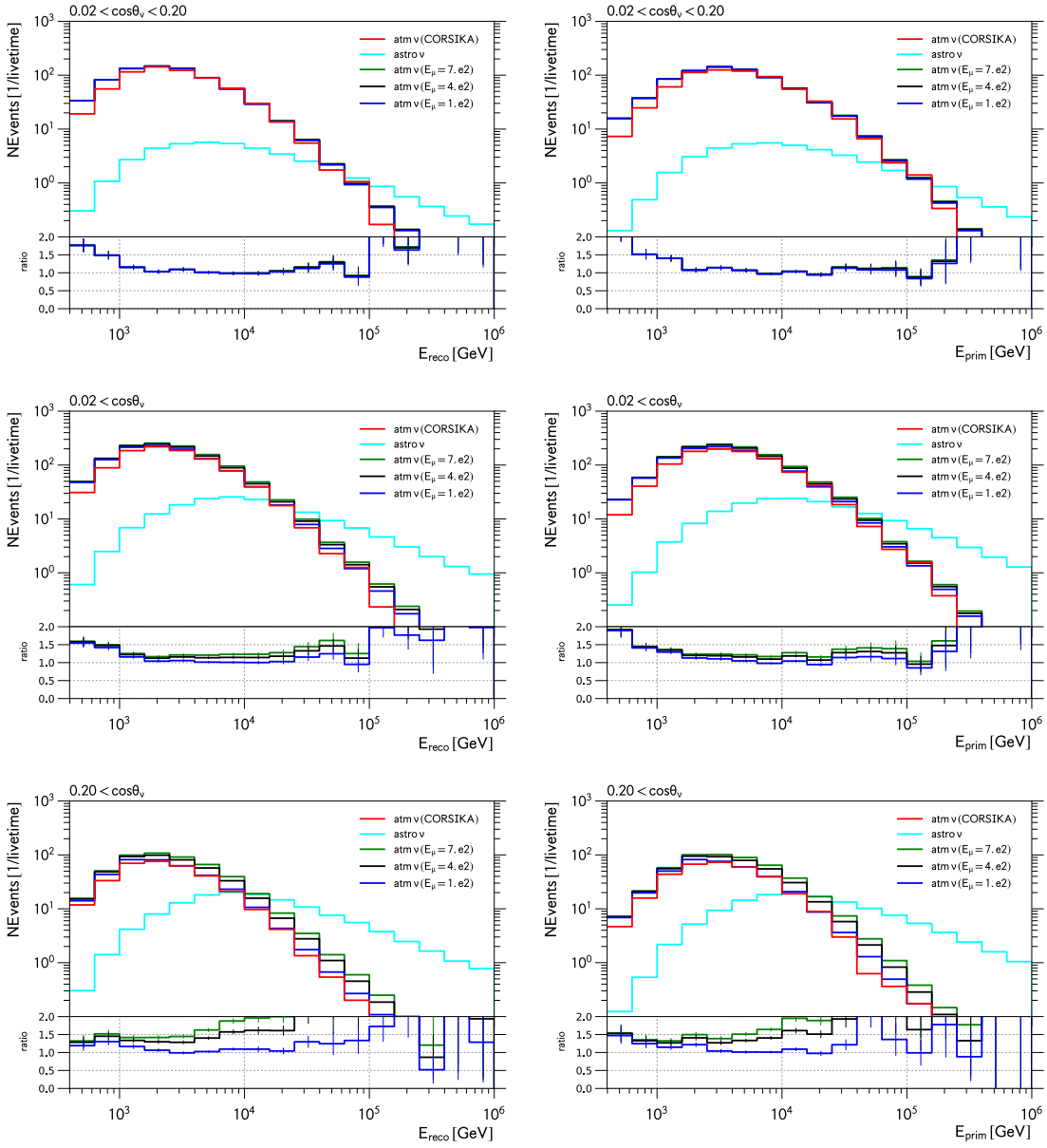


FIGURE 4.11: Atmospheric neutrino expectation from full CORSIKA simulation (red) and neutrino-only simulation with self-veto efficiency correction for different values of $E^{det}_{\mu,min}$ (blue, green, black). Reconstructed energy spectrum (left) and true neutrino energy spectrum (right). Nearly horizontal events (top), entire southern hemisphere (middle) and more vertical events (bottom).

$\cos\theta > 0.02$ since the CORSIKA simulation remains valid only up to $\theta = 89°$. Also shown is the expected astrophysical neutrino contribution (cyan) for an assumed spectrum with normalization $\Phi_{astro} = 1.6$ and spectral index $\gamma = 2.5$. The different rows in Fig. 4.11 correspond to different declination bands (top row: $0.02 < \cos\theta_\nu < 0.20$, middle row: $0.02 < \cos\theta_\nu$, bottom row: $0.20 < \cos\theta_\nu$). The top row shows events with nearly horizontal trajectories. Due to the drastic increase of the ice-overburden and thus shielding from muons towards the horizon (c.f. Fig. 4.10 (bottom)), no self-veto effect

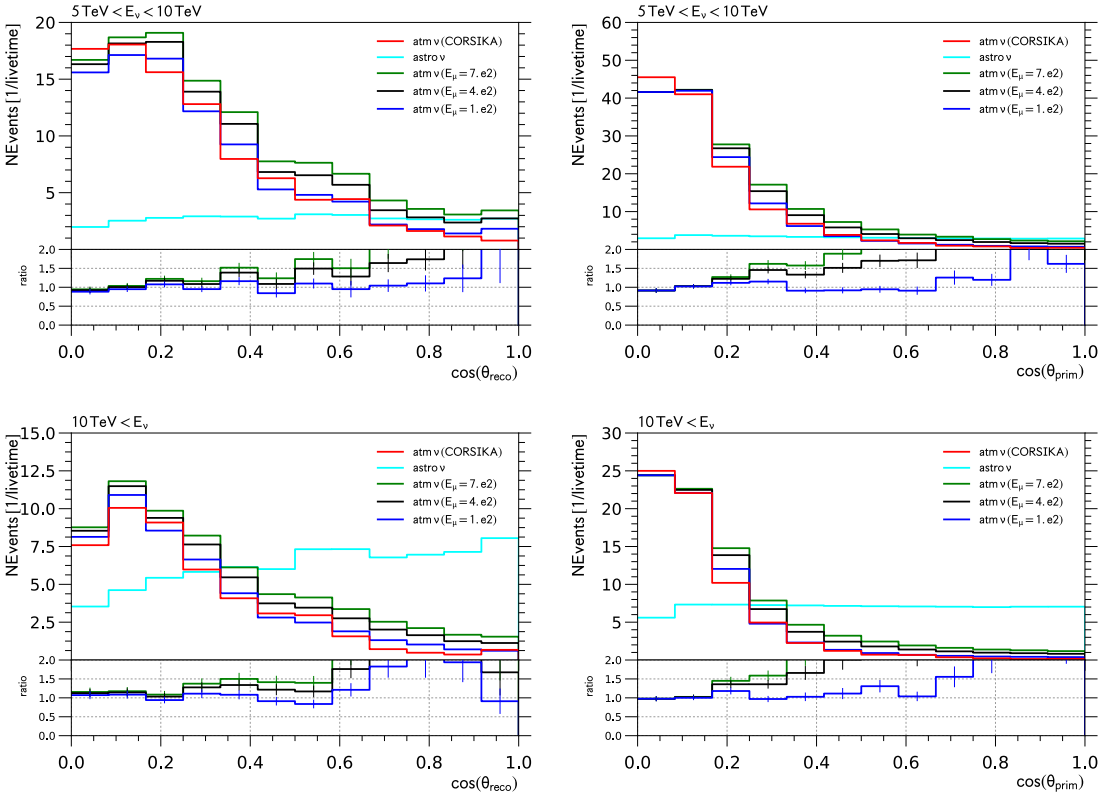


FIGURE 4.12: Zenith distributions of atmospheric neutrinos predicted from CORSIKA (red) and self-veto corrected neutrino simulations for different values of $E^{det}_{\mu,min}$ (blue, green, black). reconstructed zenith (left), true neutrino zenith (right). 5 TeV $< E_{reco} <$ 10 TeV (top) and 10 TeV $< E_{reco}$ (bottom)

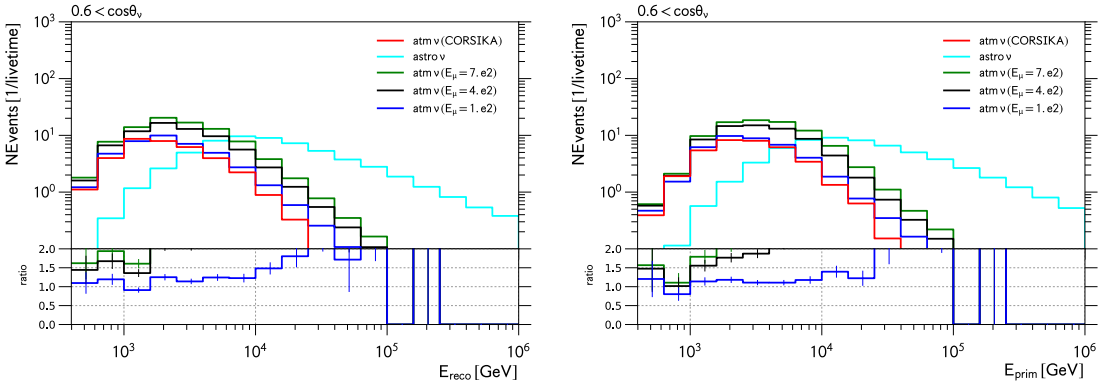


FIGURE 4.13: Atmospheric neutrino expectation from full CORSIKA simulation (red) and neutrino-only simulation with self-veto efficiency correction for different values of $E^{det}_{\mu,min}$ (blue, green, black). Reconstructed energy spectrum (left) and true neutrino energy spectrum (right). Nearly vertical directions.

is observed. We find good agreement between neutrino-only simulation (with weights determined by MCEq) and full CORSIKA simulation for most of the energy range. At

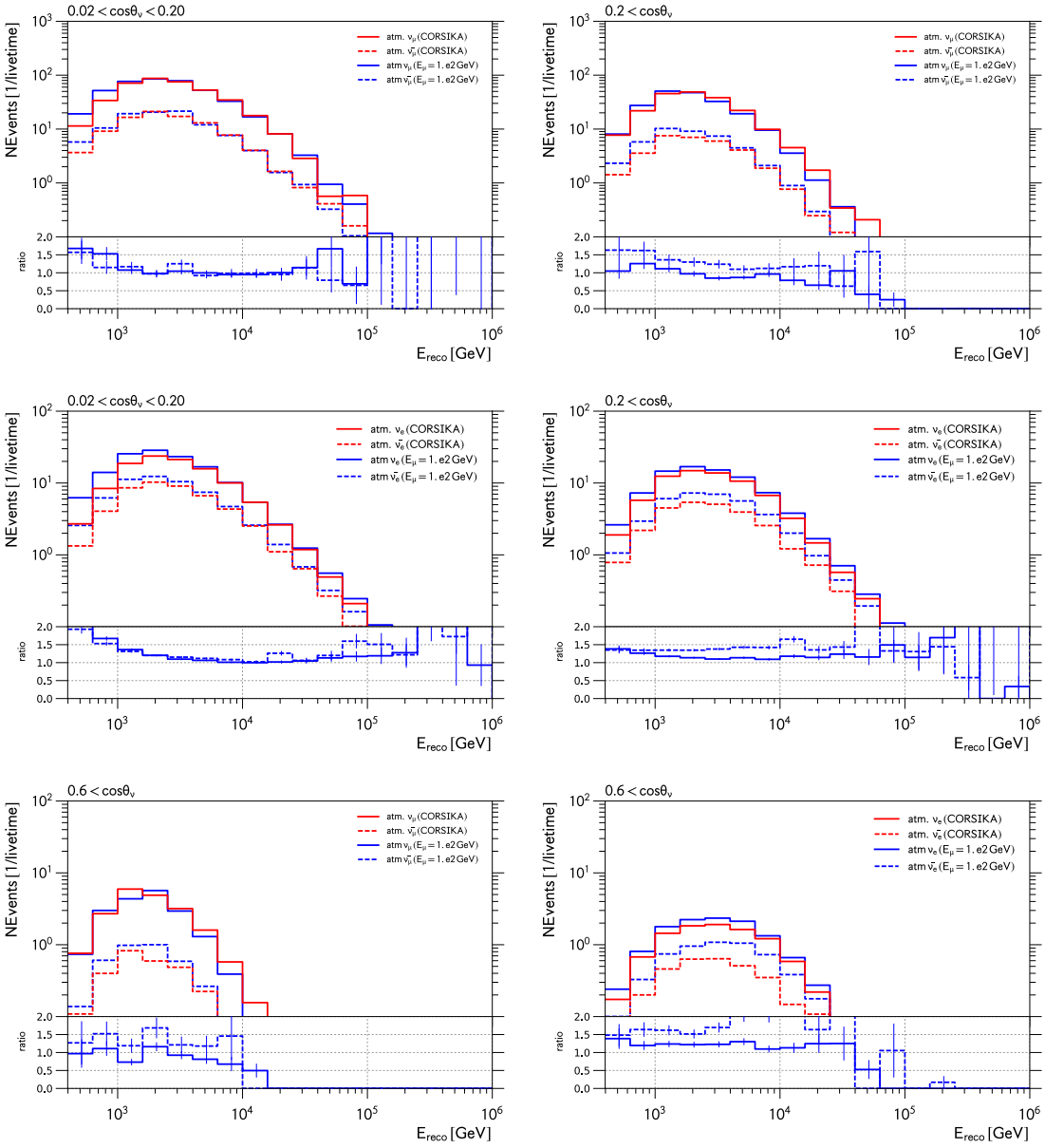


FIGURE 4.14: Atmospheric neutrino expectation from full CORSIKA simulation (red) and neutrino-only simulation with self-veto efficiency correction (blue). muon neutrinos (top), electron neutrinos (middle). $0.02 < \cos\theta_\nu < 0.20$ (left), $0.20 < \cos\theta_\nu < 1.00$ (right). Bottom: near vertical trajectories for muon neutrinos (left) and electron neutrinos (right)

neutrino energies above $\sim$ 30 TeV the comparison is limited by the total number of simulated air-showers in CORSIKA. At lowest neutrino energies (below a few TeV) systematic discrepancies become visible but are considered irrelevant for the purpose of validating the atmospheric self-veto modeling. The bottom row corresponds to events that have large probabilities of being accompanied by muons and thus are subject to significant vetoing. Reasonable agreement between the CORSIKA and neutrino-only spectra are observed for $E^{det}_{\mu,min} = 100\,\text{GeV}$, especially in the important transition region near $E_{reco} = 10\,\text{TeV}$

where the composition changes from being dominated by atmospheric neutrinos to being dominated by extra-terrestrial neutrinos. Increasing the veto threshold $E^{det}_{\mu,min}$ above 100 GeV significantly over-estimates the CORSIKA prediction (green, black).
The predicted zenith distributions are shown in Fig. 4.12 for two energy ranges $5\,\text{TeV} < E_{reco} < 10\,\text{TeV}$ (top) and $10\,\text{TeV} < E_{reco}$ (bottom). In the spectral fits presented in this dissertation not all zenith angle information is used, since only three observable bins contribute: the energy spectrum in the northern sky ($\cos\theta_{rec} < 0.2$), the energy spectrum in the southern sky for near horizontal trajectories ($0.2 \leq \cos\theta_{rec} < 0.6$) and, finally, near vertical trajectories with $\cos\theta_{rec} \geq 0.6$. However it is still interesting to study the agreement of the predict zenith distributions between both approaches. Very good agreement is observed, especially in the important energy range $5\,\text{TeV} < E_{reco} < 10\,\text{TeV}$ for both, true and reconstructed, observables. Above 10 TeV the agreement remains good, except for an excess in the prediction from CORSIKA over the neutrino-only simulation for very vertical events. It indicates a systematic under-estimation of the veto efficiency at high energies by up to a factor of two for the vertical direction, but is not relevant for the purpose of this work. In this region of observable space the atmospheric contribution to the spectrum is sufficiently subdominant compared to the contribution from astrophysical neutrinos (cyan), such that the difference falls significantly below the sensitivity of the experiment for the given total live time of four years (c.f. Fig. 4.13).
We finally separated the simulations by neutrino type, since the different neutrinos are sensitive to different aspects of the veto-modeling. Muon neutrinos will be more sensitive to the modeling of the correlations between muons and neutrinos from the same meson decay, while electron neutrinos are only affected by the modeling of the correlation between neutrino and muon stemming from different meson decays within the air-shower. The corresponding reconstructed energy distributions are shown in Fig. 4.14 (top: ν_μ, bottom: ν_e) for two zenith bands. While in general both simulations agree well, the expected contribution of electron antineutrinos $\bar{\nu}_e$ ($\cos\theta_\nu > 0.2$) from self-veto corrected neutrino simulation appears systematically larger than the expectation from CORSIKA simulation by $\sim 40\%$. Since the overall contribution from $\bar{\nu}_e$ to the spectrum is sufficiently small, this does not pose a problem for this analysis. Further studies are required to resolve the remaining discrepancies.

In summary, we demonstrated with full CORSIKA airshower simulations that the parametrization of the self-veto probabilities from [179][180] is sufficiently accurate to be used in this analysis. It allows to calculate the expected atmospheric neutrino background in the southern hemisphere from neutrino-only simulations and we find a veto-threshold of $E^{det}_{\mu,min} = 100\,\text{GeV}$ to perform best. The remaining discrepancies between this model and CORSIKA simulations appear small compared to the other systematic uncertainties discussed in this chapter and the corresponding systematic uncertainty will thus not be included in the spectral likelihood fits presented in this dissertation.

Chapter 5

Analysis Method and Statistics

In this dissertation we measure and characterize the astrophysical neutrino flux by matching the expected energy and zenith distributions in our final data sample obtained from Monte-Carlo simulations to the distributions observed in the experimental data. The simulations include all relevant physics (particle production, propagation, interaction, light propagation, response of the detector electronics etc.) and measurement (event selection) processes for different choices of astrophysical and atmospheric neutrino flux models as well as different models of the detector response.

5.1 Modeling the Astrophysical Neutrino Flux

Throughout this analysis we assume the fluxes of the three neutrino flavors to be identical ($\Phi_{\nu_e} : \Phi_{\nu_\mu} : \Phi_{\nu_\tau} = 1 : 1 : 1$), independent of neutrino energy. Similarly, we assume the contributions from neutrinos and anti-neutrinos to each flavor to be equal ($\Phi_{\nu_x} : \Phi_{\bar{\nu}_x} = 1 : 1$). Finally, the astrophysical neutrino flux is assumed independent of arrival direction and thus modeled isotropic: $\Phi\left(E,\, \cos\theta_{dec.},\, \phi_{ra.}\right) = \Phi\left(E\right)$.

5.1.1 Parametric Models

The Single Powerlaw

The default model for the astrophysical neutrino flux is the single powerlaw. It is fully specified by two parameters: the spectral index γ and a per-flavor normalization Φ_0 corresponding to the total flux of neutrinos with flavor $x \in \{e,\, \mu,\, \tau\}$ at $E = 100\,\mathrm{TeV}$ with $\Phi_0 = \Phi_{\nu_x + \bar{\nu_x}}(E/\mathrm{TeV} = 100)$:

$$\Phi\left(E\right)/(\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}) = 3 \times 10^{-18}\,\Phi_0 \times \left(\frac{E}{100\,\mathrm{TeV}}\right)^{-\gamma} \qquad (5.1)$$
$$\Phi_0 \geq 0$$

The Single Powerlaw with Cutoff

The single powerlaw model, while considered appropriate to model the neutrino flux in a limited energy range, cannot be true regardless of energy. In order to preserve energy conservation any physically motivated acceleration process cannot accelerate particles to arbitrarily high energies. We therefore extend the single power law by an additional parameter E_{cut} denoting the energy beyond which the total flux is assumed to vanish exponentially:

$$\Phi\left(E\right)/(\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}) = 3\times10^{-18}\,\Phi_0 \times \left(\frac{E}{100\,\mathrm{TeV}}\right)^{-\gamma} \times \exp\left(-\frac{E}{E_{cut}}\right) \qquad (5.2)$$
$$\Phi_0 \geq 0, \quad \frac{E_{cut}}{\mathrm{TeV}} > 0$$

The model reduces to the (unbroken) single powerlaw in the limit of $E_{cut} \to \infty$.

The 2-Component Powerlaw

Since the true origin of the high energy astrophysical neutrinos, observed by IceCube, is unknown at present, we consider the possibility that neutrinos are produced by different types of objects. Here we consider the simplest extension of the single powerlaw to two objects. We allow for the possibility that the fluxes from the different classes of objects are different in intensity and shape. Thus the model has four parameters: the spectral index γ_i and a per-flavor normalization $\Phi_{i,0}$ with $i \in \{1,2\}$. Since the likelihood function will be invariant under the label-switch $i = 1 \to 2$ and $i = 2 \to 1$ it is useful to impose an identifiability constraint [183]. Without loss of generality, we choose $\gamma_1 > \gamma_2$, i.e. the second component in this power law mixture is identified as the one that has the harder spectrum. The model becomes:

$$\Phi\left(E\right)/(\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}) = 3\times10^{-18}\sum_{i=1,2}\Phi_{i,0} \times \left(\frac{E}{100\,\mathrm{TeV}}\right)^{-\gamma_i} \qquad (5.3)$$

$$\Phi_{i,0} \geq 0, \quad \gamma_1 > \gamma_2 \qquad (5.4)$$

The model reduces to the (unbroken) single powerlaw for $\Phi_{2,0} = 0 \times \mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$. The constraint in eq. (5.4) cannot easily be integrated into numerical minimization algorithms. If numerical minimization is used, we thus use a different, but equivalent parametrization:

$$\frac{\Phi\left(E\right)}{\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}} = 3\times10^{-18}\,\Phi_0\left\{(1-\alpha)\left[\frac{E}{100\,\mathrm{TeV}}\right]^{-\gamma_{\mathrm{soft}}} + \alpha\left[\frac{E}{100\,\mathrm{TeV}}\right]^{-\gamma_{\mathrm{soft}}+\Delta\gamma}\right\}$$
$$\Phi_0 \geq 0, \quad 0 < \alpha < 1, \quad \Delta\gamma \geq 0 \qquad (5.5)$$

with $\gamma_{soft} = \gamma_1$ and $\gamma_{soft} - \Delta\gamma = \gamma_2$ where we refer to $\Delta\gamma$ as the spectral hardening. Finally Φ_0 denotes the total flux at $E = 100\,\mathrm{TeV}$ and relates to the previous parametrization

with the mixing fraction α: $\Phi_{1,0} = (1-\alpha)\,\Phi_0$ and $\Phi_{2,0} = \alpha\,\Phi_0$. In this parametrization the (unbroken) single powerlaw is recovered for $\Delta\gamma = 0$.

The Broken Powerlaw

The broken power law, in contrast to the 2-component power law, can represent both: a spectral hardening as well as a spectral steepening of the neutrino flux at some break energy E_b. Spectral breaks in the fluxes of cosmic rays have been observed for various types of particles, for example in protons [15] and electrons/positrons [184]. In addition certain models for neutrino emission from gamma-ray bursts predict spectral breaks [39]. Here we consider the simplest case of a single spectral break.

$$\frac{\Phi\,(E)}{\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}} = 3\times10^{-18}\Phi_b \times \begin{cases} (E\,/\,E_b)^{-\gamma_1} & E \le E_b \\ (E\,/\,E_b)^{-\gamma_2} & E > E_b \end{cases} \qquad E_b > 0\,\mathrm{TeV}, \quad \Phi_b > 0 \tag{5.6}$$

This model has four parameters: the break energy E_b, spectral indices γ_1 and γ_2 and the total flux at the break energy $\Phi_b = \Phi\,(E_b)$. In order to maintain comparability to the other models, we re-parametrize the model in terms of Φ_0, the flux at $E = 100\,\mathrm{TeV}$.

$$\Phi_b = \Phi_0 \times \begin{cases} (100\,\mathrm{TeV}\,/\,E_b)^{\gamma_1} \\ (100\,\mathrm{TeV}\,/\,E_b)^{\gamma_2}\,, \end{cases} \quad \Phi_0 > 0 \tag{5.7}$$

The model reduces to the single powerlaw for $\gamma_1 = \gamma_2$.

The Log-Parabolic Powerlaw

γ-ray spectra of certain types of blazars have successfully been modeled with log-parabolic powerlaws [185]. In such spectra the spectral index varies as function of energy. The same parametrization has also been used to model the proton component in blazar jets [186], producing neutrino spectra of similar shape. We use the following parametrization:

$$\begin{aligned} \frac{\Phi\,(E)}{\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}} &= 3\times10^{-18}\,\Phi_0 \times \left(\frac{E}{100\,\mathrm{TeV}}\right)^{-\Gamma(E)} \\ \Gamma\,(E) &= \gamma + b\,\log\left(E/100\,\mathrm{TeV}\right) \\ \Phi_0 \ge 0, &\quad b \ge 0 \end{aligned} \tag{5.8}$$

This model consists of three parameters: spectral index γ at $E = 100\,\mathrm{TeV}$, total flux Φ_0 at $E = 100\,\mathrm{TeV}$ and the log-parabolic slope b that controls the logarithmic energy dependence of γ. The model recovers the single powerlaw for $b = 0$.

Fig. 5.1 compares the compares the different parametric models discussed above, starting from the single powerlaw (black) with $\gamma = 2.5$ and $\Phi_0 = 1.6$.

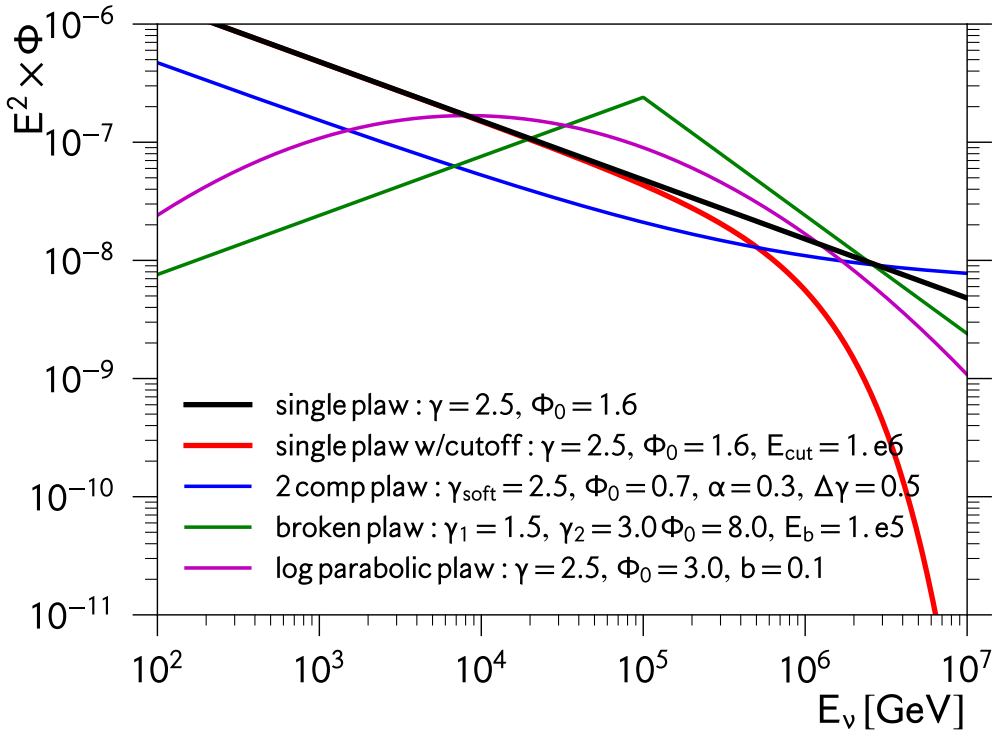


FIGURE 5.1: Various parametric astrophysical neutrino flux models: the single power-law (black) and extensions thereof (colors).

The Single Powerlaw Box

The single powerlaw assumes that the astrophysical neutrino flux is non-zero at all energies. Here we expand this model by limiting its domain. While not motivated by any physics mechanism this model will be useful in the context of sensitivity studies by addressing the question what energy range the measurement of the astrophysical neutrino flux presented in this dissertation is sensitive to. The "box" is characterized by two additional parameters: E_{low} and E_{high} that define the domain where the astrophysical flux is non-zero

$$\Phi(E) = \Phi_{SP}(E) \times I(E) \tag{5.9}$$

$$I(E) = \begin{cases} 1 & E_{low} \leq E \leq E_{high} \\ 0 & \text{otherwise} \end{cases} \tag{5.10}$$

where Φ_{SP} denotes the single power law with infinite domain (entire real line). Of course any of the models discussed above could be "boxified" similarly.

The 2-Hemisphere Model

Thus far the different flux parameterizations, discussed above, assume the astrophysical neutrino flux to be isotropic. This assumption can be relaxed by allowing it to differ

between northern and southern hemispheres. The simplest possibility would be different powerlaws in both hemispheres.

$$\frac{\Phi\left(E\right)}{\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}} = 3\times10^{-18}\times\begin{cases}\Phi_N\times\left(E\,/\,100\,\mathrm{TeV}\right)^{-\gamma_N} & \cos\theta_\nu < 0.0\\ \Phi_S\times\left(E\,/\,100\,\mathrm{TeV}\right)^{-\gamma_S} & \cos\theta_\nu \geq 0.0\end{cases} \qquad (5.11)$$
$$\Phi_N \geq 0, \quad \Phi_S \geq 0$$

While not necessarily corresponding to a scenario that is motivated by concrete astrophysics, this model allows to test to what extent some of the conclusions drawn in this work depend on the assumption of isotropy.

5.1.2 Non-Parametric Models

It is possible that none of the parametric models introduced above adequately describes the true astrophysical neutrino flux. It is therefore helpful to study to what extent the data can constrain the flux if less assumptions about its intensity and shape are made[1]. This can be done by introducing a class of more flexible models that admit local changes of the flux as function of energy. This increases variance in the resulting function estimate compared to using one of the parametric models but would reduce bias if the true flux is not well represented by any of the parametric models.

The Piecewise Model

We define the "Piecewise Model" as the following step-function with N steps

$$\left(\frac{E}{1\mathrm{GeV}}\right)^2\times\Phi\left(E\right) = 3\times10^{-8}\sum_{i=1}^{N}\left\{\Phi_i\times I_i\left(E\right)\right\} \qquad (5.12)$$

$$I_i\left(E\right) = \begin{cases}1 & E_i^{low}\leq E\leq E_i^{high}\\ 0 & \text{otherwise}\end{cases}, \quad \Phi_i\geq 0 \qquad (5.13)$$

It models the astrophysical neutrino flux as a segmented powerlaw with spectral index $\gamma = 2.0$ and variable normalization. The flexibility of the model increases with larger values of N. If all normalizations are identical (i.e. $\Phi_i = \Phi_j,\, \forall\, i,j : 1\leq i,j\leq N$) the solution is identical to the single powerlaw with $\gamma = 2.0$. This model requires certain choices: the total number of segments N and the boundaries E_i of each segment: $N = 12$ and $\log_{10}\left(E_i^{min}/\mathrm{GeV}\right) = 3.0 + (j-1)/3$ and $\log_{10}\left(E_i^{max}\right) = \log_{10}\left(E_i^{min}\right) + 1/3$, corresponding to three segments per decade in neutrino energy.

[1] The non-parametric models used in this work remain finite dimensional. Our interpretation of the term non-parametric thus follows the usage in [187].

5.2 Modeling the Atmospheric Neutrino Flux

5.2.1 Conventional Atmospheric Neutrinos

Following previous IceCube measurements (e.g. [5][164][181][182][155]) we use the HKKMS06 flux model [99] to characterize the contribution from conventional atmospheric neutrinos to the data sample obtained in this dissertation (c.f Sec. 1.7).

As discussed in Sec. 4.2.1 other models exist and their impact on the astrophysical neutrino flux measurement needs to be evaluated. For this purpose we rely on MCEq [169] to calculate the predicted conventional atmospheric neutrino flux for the different hadronic interaction models:

- Primary Cosmic Ray Flux Model: Gaisser-Hillas 2012, "H3a" [170]
- Atmospheric Density Model: NRLMSISE-00 [188]
- Hadronic Interaction Models:
  SIBYLL 2.3c[173], QGSJETII-04[174], EPOS-LHC[175] and DPMJETIII-17.1[176][177]

We are only interested in the flux expectation averaged over an entire year. We therefore perform the MCEq flux calculation for the two seasonal extremes of the atmospheric conditions (January and June) and average the resulting fluxes. Thus in total we will study five different atmospheric neutrino flux models. For each model the absolute flux normalization is treated as a free parameter.

5.2.2 Prompt Atmospheric Neutrinos

Previous IceCube measurements (e.g. [5][164][181][182][155]) used the ERS flux calculation [101] as the default model to predict the contribution from prompt atmospheric neutrinos. Since then, the authors have revised their calculation taking into account newly available charm cross section measurements at LHC and RHIC to constrain their perturbative QCD calculations. The resulting prediction, BERSS [102] is used as the baseline model in this work. The absolute flux normalization is treated as a free parameter.

5.3 The Likelihood Function

The following exposition of the statistical methods used in this analysis will adopt the notation and definitions of [189].

After observing a random sample $\boldsymbol{X} = (X_1, ..., X_n)$ from the probability density $f(x \,|\, \boldsymbol{\theta})$ ($x \in \Omega$, $\boldsymbol{\theta} \in \Theta$) with realized values $\boldsymbol{X} = \mathbf{x}$, we define the likelihood function $L(\boldsymbol{\theta} \,|\, \mathbf{x})$ as follows

$$L(\boldsymbol{\theta} \,|\, \mathbf{x}) = f(\mathbf{x} \,|\, \boldsymbol{\theta}) \tag{5.14}$$

where $f(\mathbf{x} \mid \boldsymbol{\theta})$ is the joint probability density of the random sample $\boldsymbol{X} = \mathbf{x}$. Eq. (5.14) implies that $L(\boldsymbol{\theta} \mid \mathbf{x})$ and $f(\mathbf{x} \mid \boldsymbol{\theta})$ are very different functions. The domain of $f(\mathbf{x} \mid \boldsymbol{\theta})$ is the entire joint sample space $\boldsymbol{\Omega} = \Omega^n$ while the domain of $L(\boldsymbol{\theta} \mid \mathbf{x})$ is the parameter space $\boldsymbol{\Theta}$. In particular $L(\boldsymbol{\theta} \mid \mathbf{x})$ is not normalized on $\boldsymbol{\Theta}$ while $f(\mathbf{x} \mid \boldsymbol{\theta})$, being a probability density, is normalized on $\boldsymbol{\Omega}$. Discussing the importance of the likelihood function for statistical inference goes beyond the scope of this dissertation and we refer to [189][190] and references therein.

In this analysis one would ideally define the sample as the set of all observed events and their associated observables (reconstructed cascade variables: energy, direction, position, and others, e.g. BDT scores) and construct the likelihood function from the joint (high dimensional) pdf according to eq. (5.14). While this would make maximal use of the available information, the resulting likelihood function would be intractable. Instead we simplify the problem by focusing on the observables that provide the most information about the energy spectrum as well as on whether or not events can be considered to be of astrophysical origin. These are the reconstructed energy, the reconstructed zenith angle and the three BDT scores: cascade, muon and starting track scores. Evaluating their joint pdf remains intractable. We will therefore sort the events according to their observables into bins, thus we discretize the observable space. We will define and justify the binning used in this analysis later (see Sec. 5.6).
The discretized problem can be viewed as individual counting experiments in each bin i of the analysis with total number of bins m. Thus $\boldsymbol{X} = \boldsymbol{N}$ and $N_i \sim \text{Poisson}(\mu_i(\boldsymbol{\theta}))$. After observing $\boldsymbol{N} = (n_1, ..., n_m)$ we arrive at the following likelihood function using eq. (5.14)

$$L(\boldsymbol{\theta} \mid \mathbf{n}) = f(\mathbf{n} \mid \boldsymbol{\mu}(\boldsymbol{\theta})) = \Pi_{i=1}^{m} p(n_i \mid \mu_i(\boldsymbol{\theta})) \tag{5.15}$$

$$p(n_i \mid \mu_i(\boldsymbol{\theta})) = \frac{\mu_i(\boldsymbol{\theta})^{n_i}}{n_i!} \exp(-\mu_i(\boldsymbol{\theta})) \tag{5.16}$$

For reasons of numerical stability it is beneficial to work with the log-likelihood

$$\log L(\boldsymbol{\theta} \mid \mathbf{n}) = \sum_{i=1}^{m} \{n_i \log \mu_i(\boldsymbol{\theta}) - \mu_i(\boldsymbol{\theta})\} + g(\boldsymbol{n}) \tag{5.17}$$

where $g(\boldsymbol{n})$ consists of all terms that are independent of $\boldsymbol{\theta}$ and thus contain no information about the relative merits of two hypothesis $\boldsymbol{\theta} = \boldsymbol{\theta_1}$ and $\boldsymbol{\theta} = \boldsymbol{\theta_2}$. Therefore $g(\boldsymbol{n})$ can be ignored.

Including Calibration Measurements: Penalized Likelihood

It is often the case that additional and independent observations $\mathbf{y}$ from a density $g(y \mid \omega)$ are available, where ω denotes one of the components of the parameter vector $\boldsymbol{\theta}$. For example consider ω as an overall scaling factor that controls the amount of bulk ice scattering in the ice. In this case $\mathbf{y}$ refers to an independent calibration dataset obtained from IceCube's LED calibration system.

Using the independence of $\mathbf{n}$ and $\mathbf{y}$ one arrives at the joint log-likelihood

$$\log L\left(\boldsymbol{\theta}_{-\omega},\, \omega \,|\, \mathbf{n}, \mathbf{y}\right) = \log L\left(\boldsymbol{\theta}_{-\omega},\, \omega \,|\, \mathbf{n}\right) + \log L\left(\omega \,|\, \mathbf{y}\right) \tag{5.18}$$

where we made the dependence on ω explicit by removing it from the parameter vector $\boldsymbol{\theta}$ thus arriving at the reduced parameter vector $\boldsymbol{\theta}_{-\omega}$. In practice a joint analysis quickly becomes intractable, especially if several calibration measurements are considered. The problem can be simplified by using a known parametrization (if available) of $\log L\left(\omega \,|\, \mathbf{y}\right)$. From asymptotic considerations (see Sec. 5.4, especially Eq. (5.23)) we know that $\hat{\omega} - \omega \xrightarrow{d} N\left(0 \,|\, \sigma^2_{\hat{\theta}}|_{n\to\infty}\right)$. Thus, assuming that asymptotic theory applies, we can ignore $\mathbf{y}$ and write the contribution to the log-likelihood function in terms of the measured values $\hat{\omega}$ and $\sigma_{\hat{\omega}}$ from the calibration measurement (based on data $\mathbf{y}$):

$$\log L\left(\omega \,|\, \mathbf{y}\right) = \log L\left(\omega \,|\, \hat{\omega}, \sigma_{\hat{\omega}}\right) = -\frac{\left(\hat{\omega} - \omega\right)^2}{2\sigma^2_{\hat{\omega}}} \tag{5.19}$$

where we dropped terms independent of ω. Thus the final joint log-likelihood reads:

$$\log L\left(\boldsymbol{\theta}_{-\boldsymbol{\omega}},\, \boldsymbol{\omega} \,|\, \mathbf{n}, \hat{\boldsymbol{\omega}}, \boldsymbol{\Sigma}\right) = \left\{\sum_{i=1}^{m} n_i \log \mu_i\left(\boldsymbol{\theta}\right) - \mu_i\left(\boldsymbol{\theta}\right)\right\} - \frac{1}{2}\left(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}\right)^T \boldsymbol{\Sigma}^{-1}\left(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}\right) \tag{5.20}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix allowing for correlations between parameters of the calibration measurements and generalizes Eq. (5.19) to the multivariate case.

Throughout this work we will use the penalized likelihood from Eq. (5.20).

5.4 Point Estimation, Hypothesis Testing and Interval Estimation

Point Estimation

We will estimate unknown parameters $\boldsymbol{\theta}$ of a model using the *method of maximum likelihood* unless mentioned otherwise. The maximum likelihood estimator $\hat{\boldsymbol{\theta}}\left(\mathbf{x}\right)$ is defined as the point $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}\left(\mathbf{x}\right)$ $\left(\boldsymbol{\theta} \in \boldsymbol{\Theta}\right)$ at which the likelihood function (Eq. (5.14)) attains is maximum value (with $\mathbf{x}$ kept fixed at its observed value).

$$\hat{\boldsymbol{\theta}}\left(\mathbf{x}\right) = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmax}}\; L\left(\boldsymbol{\theta} \,|\, \mathbf{x}\right) \tag{5.21}$$

This is equivalent to minimizing the negative log-likelihood, which is preferred for numerical reasons.

$$\hat{\boldsymbol{\theta}}\left(\mathbf{x}\right) = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmin}}\; \left(-\log\, L\left(\boldsymbol{\theta} \,|\, \mathbf{x}\right)\right) \tag{5.22}$$

Maximum likelihood estimators (MLEs) have several attractive properties. They are invariant under parameter transformations and (under some regularity conditions) consistent[2] and asymptotically efficient[3]. Finally the asymptotic variance of the MLE is related to the fisher information matrix $I(\theta)$ (under some regularity conditions)

$$\sigma^2_{\hat{\theta}}|_{n\to\infty} = \operatorname{Var}\hat{\theta}|_{n\to\infty} = \frac{1}{I(\theta)} = -\frac{1}{E_\theta\left(\frac{\partial^2}{\partial\theta^2}\log L(\theta\,|\,\mathbf{X})\right)} \tag{5.23}$$

or in multivariate form with covariance matrix $\Sigma_{i,j} = (I(\boldsymbol{\theta}))^{-1}_{i,j}$:

$$I(\boldsymbol{\theta})_{i,j} = -E_{\boldsymbol{\theta}}\left(\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\log L(\boldsymbol{\theta}\,|\,\mathbf{X})\right) \tag{5.24}$$

Eq. (5.24) will be useful in the context of calculation covariance matrices from previous measurements. For more details and definitions see [189].

Hypothesis Testing

Unless explicitly stated otherwise, throughout this thesis we will rely on *likelihood ratio tests* (LRTs) to distinguish based on the observed data between two complementary hypotheses, the *null hypothesis* H_0 and the *alternative hypothesis* H_1. The test is based on the *test statistic* $\Lambda(\mathbf{x})$ defined as follows

$$\Lambda(\mathbf{x}) = \frac{\sup_{\boldsymbol{\Theta_0}} L(\boldsymbol{\theta}\,|\,\mathbf{x})}{\sup_{\boldsymbol{\Theta}} L(\boldsymbol{\theta}\,|\,\mathbf{x})} \tag{5.25}$$

for testing the hypothesis $H_0 : \boldsymbol{\theta} \in \boldsymbol{\Theta_0}$ against the alternative $H_1 : \boldsymbol{\theta} \in \boldsymbol{\Theta_0^c}$. The test rejects H_0 for small values of $\Lambda(\mathbf{x})$, i.e. its rejection region has the form $\{\mathbf{x} : \Lambda(\mathbf{x}) \leq c\}$, with $0 \leq c \leq 1$ being a constant.
For the most part, hypothesis tests employed in this thesis are *two-sided*, i.e. of form $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. In some cases we will test *one-sided* hypothesis of the form $H_0 : \boldsymbol{\theta} \geq \boldsymbol{\theta}_0$ ($H_0 : \boldsymbol{\theta} \leq \boldsymbol{\theta}_0$) against $H_1 : \boldsymbol{\theta} < \boldsymbol{\theta}_0$ ($H_1 : \boldsymbol{\theta} > \boldsymbol{\theta}_0$). In order to find the constant c that bounds the rejection region of the test one can specify a desired value for the *size* α ($0 \leq \alpha \leq 1$) of the test, where α is defined as the maximum Type-1 error probability across H_0 ($\boldsymbol{\theta} \in \boldsymbol{\Theta_0}$). The corresponding value $c = c(\alpha)$ can then be determined from the *sampling distribution* of $\Lambda(\mathbf{x})$ (see [189] for details). In general, depending on the problem, finding c can be hard if not numerically impossible, especially if the problem contains several *nuisance parameters*. These are parameters of the model that do not explicitly enter the hypotheses being tested. In the context of this dissertation these are those parameters that are unrelated to the astrophysical neutrino flux, e.g. those that are needed to model the background fluxes and the detector response.

[2] The MLE converges to the true value as the sample size increases

[3] The variance of the MLE converges to Cramer Rao lower bound, a lower bound on the variance of all possible estimators, as the sample size increases.

For large samples sizes $c(\alpha)$ can be obtained approximately by invoking *Wilks' theorem* [191]. Assuming certain regularity conditions the sampling distribution of $-2\log\Lambda(\mathbf{x})$, with $\Lambda(\mathbf{x})$, as defined in Eq. (5.25), approaches a chi-squared distribution if $H_0 : \boldsymbol{\theta} \in \boldsymbol{\Theta_0}$ is true:

$$-2\log\Lambda(\mathbf{x}) \xrightarrow{d} \chi^2_k \tag{5.26}$$

where $k = \dim(\Theta) - \dim(\Theta_0)$, i.e the degrees of freedom are given by the difference in the number of free parameters between both hypotheses. We confirmed the applicability of Wilk's theorem in the context of this analysis and the single-powerlaw model for hypothesis tests of form $H_0 : (\Phi, \gamma) = (\Phi_0, \gamma_0)$ against $H_1 : (\Phi, \gamma) \neq (\Phi_0, \gamma_0)$ by calculating the expected sampling distribution of $-2\log\Lambda(\mathbf{x})$ from repeatedly performing the test on replicated toy-data. The result is shown in Fig. 5.2 and matches the expectation from eq. (5.26) very well.

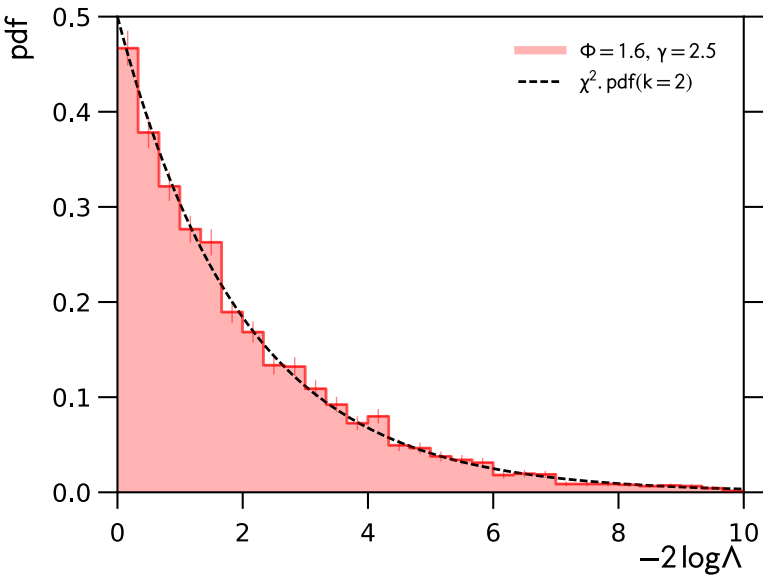


FIGURE 5.2: Distribution of the profile-likelihood ratio test-statistic $-2\log\Lambda$ used to define confidence regions: calculated from simulated toy experiments (red) and compared to asymptotic expectation from Wilks' thereom [191] (black).

For uni-variate one-sided tests one can show that the sampling distribution is of mixed type, with $p(-2\log\Lambda(\mathbf{x}) = 0) = 1/2$ and the remaining probability being distributed as χ^2_1 [192][193]. Define $\xi(\mathbf{x}) = -2\log\Lambda(\mathbf{x})$ then

$$\xi(\mathbf{x}) \xrightarrow{d} f_\xi(\xi) \tag{5.27}$$

$$f_\xi(\xi) = \frac{1}{2} \times I(0) + \chi^2_1(\xi) \tag{5.28}$$

$$I(\xi) = \begin{cases} 1 & \xi = 0 \\ 0 & \xi > 0 \end{cases} \tag{5.29}$$

In cases where asymptotic considerations are inappropriate or regularity conditions are violated we will determine the sampling distribution $f_\xi(\xi \,|\, \boldsymbol{\theta})$ from simulating repeated

measurements. Finally the corresponding p-value is defined as the probability of observing a test statistic value providing equal or more extreme evidence against the null hypothesis than the observed value ξ_{obs}[4].

$$p_{val} = \sup_{\boldsymbol{\theta}} p\left(\xi \geq \xi_{obs} \,|\, \boldsymbol{\theta}\right) \tag{5.30}$$

The rejection criterium $\xi \geq \tilde{c}$ is equivalent to $p_{val} < \alpha$.

Interval Estimation

The goal of constructing a confidence intervals or regions is to express uncertainty about the parameters of interest taking into account the measurement process $f_{\mathbf{X}}\left(\mathbf{x} \,|\, \boldsymbol{\theta}\right)$ and the observed data $\mathbf{X} = \mathbf{x}$. We define the *coverage probability* $1-\alpha$ of a $1-\alpha$ confidence region $S\left(\mathbf{X}\right)$ as the probability that the (random) region $S\left(\mathbf{X}\right)$ contains $\boldsymbol{\theta}$, i.e

$$p\left(\boldsymbol{\theta} \in S\left(\mathbf{X}\right) \,|\, \boldsymbol{\theta}\right) = 1-\alpha, \quad \forall \boldsymbol{\theta} \in \boldsymbol{\Theta} \tag{5.31}$$

In complicated problems exact coverage as defined by Eq. (5.31) cannot be guaranteed, e.g. when the coverage probability is a function of the nuisance parameters ψ. In this case we aim to guarantee minimal coverage

$$p\left(\boldsymbol{\theta} \in S\left(\mathbf{X}\right) \,|\, \boldsymbol{\theta}, \boldsymbol{\psi}\right) \geq 1-\alpha, \quad \forall \left(\boldsymbol{\theta}, \boldsymbol{\psi}\right) \in \boldsymbol{\Theta} \times \boldsymbol{\Psi} \tag{5.32}$$

Throughout this work we will obtain confidence intervals from *inverting a likelihood ratio test*[5] of size α. If, for each $\boldsymbol{\theta_0} \in \boldsymbol{\Theta}$, we let $A\left(\boldsymbol{\theta_0}\right)$ be the acceptance region (in sample space Ω) of a likelihood ratio test $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta_0}$ of size α we can construct a $1-\alpha$ confidence region from

$$S\left(\mathbf{X}\right) = \left\{\boldsymbol{\theta_0} : \mathbf{x} \in A\left(\boldsymbol{\theta_0}\right)\right\}, \quad \boldsymbol{\theta_0} \in \boldsymbol{\Theta} \tag{5.33}$$

This definition/construction is general in that in encompasses several methods of interval construction commonly used in physics [195]: *profile likelihood intervals*, *Feldman-Cousins intervals* [196] and extensions thereof (e.g. Feldman-Cousins with nuisance parameters [197]). Finally, it is worth pointing out that the shape of the confidence interval is often related to the alternative hypothesis of the LRT that is being inverted. Two-sided alternative hypotheses $H_1 : \theta \neq \theta_0$ usually produce intervals ($S\left(\mathbf{X}\right) = \left[L\left(\mathbf{X}\right), U\left(\mathbf{X}\right)\right]$) that are bounded from above and below, while the inversion of one-sided tests $H_1 : \theta < \theta_0$ ($H_1 : \theta > \theta_0$) allows to set upper (lower) limits on θ. However on bounded parameter spaces two-sided intervals may also result in limits, as exploited by Feldman and Cousins [196] for physical bounds of type $\theta \geq 0$.

When applicable we will use asymptotic profile likelihood confidence intervals and regions throughout this work. They are based on the asymptotic properties Eq. (5.26) of the likelihood ratio test-statistic of Eq. (5.25). Define $\xi\left(\boldsymbol{\theta_0} \,|\, \mathbf{x}\right) = -2\log\Lambda\left(\mathbf{x}\right)|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}}$, i.e. the

[4] alternative p-value definitions exist: see the excellent review [194]

[5] more precisely we are inverting a family of hypothesis tests, one for each value $\theta = \theta_0, \quad \theta \in \Theta$

test statistic value for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta_0}$ against $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta_0}$. Then

$$\xi\left(\boldsymbol{\theta_0} \,|\, \mathbf{x}\right) = -2\left(\log L\left(\boldsymbol{\theta_0} \,|\, \mathbf{x}\right) - \log L\left(\hat{\boldsymbol{\theta}} \,|\, \mathbf{x}\right)\right) \tag{5.34}$$

Inverting the test and using Eq. (5.26) (i.e. $\xi \xrightarrow{d} \chi^2_{dim\boldsymbol{\theta}}$) we find the following $1-\alpha$ confidence region

$$S\left(\mathbf{X} = \mathbf{x}\right) = \{\boldsymbol{\theta_0} : \xi\left(\boldsymbol{\theta_0} \,|\, \mathbf{x}\right) <= \xi_{crit}\} \tag{5.35}$$

$$\xi_{crit} = F^{-1}_{\chi^2_{dim\boldsymbol{\theta}}}\left(1-\alpha\right) \tag{5.36}$$

where $F^{-1}_{\chi^2_{dim\boldsymbol{\theta}}}$ denotes the inverse cumulative chi-squared distribution with $k = dim\,\boldsymbol{\theta}$. For $k = 1$ this yields $\xi_{crit} = 1$ corresponding to the well known interval defined by the "rule": $-2\Delta \log L \leq 1$. We intentionally left out possible nuisance parameters $\boldsymbol{\psi}$ to maintain clean notation. The results remain fully general, as Eq. (5.25) makes no choice with regard to which parts of the parameter space are considered of interest and which are considered nuisance. The usually procedure of conditional likelihood minimizations w.r.t $\boldsymbol{\psi}$ during "profile likelihood scans" of the parameters of interest $\boldsymbol{\theta}$ follows directly.

5.5 Goodness of Fit Considerations

In the previous paragraphs we discussed the central role of the likelihood function in the context of addressing questions about parameters $\boldsymbol{\theta}$ of a model M_0 assumed to be true. *Goodness of Fit Tests* are intended as a tool to question that assumption. Here we will only discuss the *saturated Poisson likelihood test* [198] employed in this work. It uses that the data is analyzed in bins corresponding to observations $\mathbf{n}$ with known distribution $n_i \sim \text{poisson}\left(\mu_i\right)$, see eqs. (5.15)-(5.18). We wish to test $H_0 : M_0\, is\, true$ against $H_1 : M_0\, is\, false$ without actually specifying an alternative model M_1. Unfortunately H_1, when formulated like that, is not a testable hypothesis. Instead we choose H_1 sufficiently flexible, such that it encompasses the set of all possible models M_i that could be constrained by the available data $\mathbf{n}$, i.e. $dim\,\boldsymbol{\Theta_i} \geq dim\,\mathbf{n} \equiv N$. The test becomes: $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}\left(\boldsymbol{\theta}\right)$ against $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}\left(\boldsymbol{\theta}\right)$, where $\boldsymbol{\mu} = \{\mu_1, ...\mu_N\}$ is the vector of the expected bin contents and $\boldsymbol{\theta}$ is the vector of parameters defined under M_0. Using eqs. (5.25) and defining $\xi\left(\mathbf{n}\right) = -2\log\Lambda\left(\mathbf{n}\right)$ we find the test-statistic advocated in [198]

$$\xi\left(\mathbf{n}\right) = 2\sum_{i=1}^{N}\left\{\mu_i\left(\hat{\boldsymbol{\theta}}\right) - n_i + n_i \log\left(\frac{n_i}{\mu_i(\hat{\boldsymbol{\theta}})}\right)\right\} \tag{5.37}$$

where $\hat{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$ as defined in eq. (5.21). Depending on the details of the mapping $\boldsymbol{\mu}\left(\boldsymbol{\theta}\right)$ one expects asymptotically (under the regularity conditions required by Wilks' theorem)

$$\xi\left(\mathbf{n}\right) \xrightarrow{d} \chi^2_k, \quad k = \dim\,\boldsymbol{\mu} - \dim\,\boldsymbol{\theta} \tag{5.38}$$

Note, these derivations are based on the un-penalized likelihood from eq. (5.18) without reference to any auxiliary calibration datasets $\mathbf{y}$ that would give rise to eq. (5.20). This

raises the question of how these should be treated.
As discussed in Sec. 5.3 the "penalty terms" in eq. (5.20) correspond to the following sampling model for the MLE of the auxiliary measurement, replacing the true variance $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\omega}}}$ with its estimate $\tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\omega}}}$

$$\hat{\boldsymbol{\omega}} \sim N\left(\boldsymbol{\omega}, \tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\omega}}}\right) \tag{5.39}$$

Using the notation from Sec. 5.3 the test becomes: $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}\left(\boldsymbol{\theta}_{-\boldsymbol{\omega}}, \boldsymbol{\omega}\right)$ against $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}\left(\boldsymbol{\theta}_{-\boldsymbol{\omega}}, \boldsymbol{\omega}\right)$, where $\boldsymbol{\omega}$ denotes the vector of nuisance parameters, as before. Superficially, this formulation looks identical to the previous one. The difference is implicit, since the test is now defined on multiple samples and thus requires the joint likelihood defined in eq. (5.20). H_0 implicitly further requires that the nuisance parameters $\boldsymbol{\omega}$ be "shared" among datasets $\mathbf{n}$ and $\mathbf{y}$, i.e. $\boldsymbol{\omega}_{\mathbf{n}} = \boldsymbol{\omega}_{\mathbf{y}} \equiv \boldsymbol{\omega}$, while H_1 does not (obviously). Thus, the dimensionality of the alternative hypothesis increased: $\dim \boldsymbol{\mu} + \dim \boldsymbol{\omega} > \dim \boldsymbol{\mu}$. After maximizing the likelihood under both hypotheses, the test-statistic[6] becomes

$$\begin{aligned}\xi\left(\mathbf{n}, \hat{\boldsymbol{\omega}}\left(\mathbf{y}\right), \tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\omega}}}\left(\mathbf{y}\right)\right) =& 2\sum_{i=1}^{N}\left\{\mu_i\left(\hat{\boldsymbol{\theta}}\right) - n_i + n_i \log\left(\frac{n_i}{\mu_i(\hat{\boldsymbol{\theta}})}\right)\right\} \\ &+ \left[\hat{\boldsymbol{\omega}}\left(\mathbf{y}\right) - \hat{\boldsymbol{\omega}}\left(\mathbf{n}, \mathbf{y}\right)\right]^T \tilde{\boldsymbol{\Sigma}}^{-1}\left[\hat{\boldsymbol{\omega}}\left(\mathbf{y}\right) - \hat{\boldsymbol{\omega}}\left(\mathbf{n}, \mathbf{y}\right)\right]\end{aligned} \tag{5.40}$$

where $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}\left(\mathbf{n}, \mathbf{y}\right) = \left(\hat{\boldsymbol{\theta}}_{-\boldsymbol{\omega}}\left(\mathbf{n}, \mathbf{y}\right), \hat{\boldsymbol{\omega}}\left(\mathbf{n}, \mathbf{y}\right)\right)$.
Note that $\xi(\mathbf{n}, \hat{\boldsymbol{\omega}}\left(\mathbf{y}\right), \tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\omega}}}\left(\mathbf{y}\right)) \geq \xi\left(\mathbf{n}\right)$ for any (fixed) observation $\mathbf{n}$, $\mathbf{y}$ and model M with parameters $\boldsymbol{\theta}$, since the calibration datasets $\mathbf{y}$ effectively constrain the Models ability to adept to the data $\mathbf{n}$.
Depending on the details of the mapping $\boldsymbol{\mu}\left(\boldsymbol{\theta}\right)$ one expects asymptotically (under the regularity conditions required by Wilks' theorem)

$$\xi\left(\mathbf{n}, \hat{\boldsymbol{\omega}}\left(\mathbf{y}\right), \tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\omega}}}\left(\mathbf{y}\right)\right) \xrightarrow{d} \chi^2_k, \quad k = \dim\, \boldsymbol{\mu} - \dim\, \boldsymbol{\theta}_{-\boldsymbol{\omega}} \tag{5.41}$$

which means $E[\xi(\mathbf{n}, \hat{\boldsymbol{\omega}}\left(\mathbf{y}\right), \tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\omega}}}\left(\mathbf{y}\right))] > E[\xi\left(\mathbf{n}\right)]$[7], as discussed above. We do not anticipate the data sample obtained in this work to satisfy asymptotic requirements, since several low statistics bins are expected/observed. We will thus determine the sampling distribution of $\xi(\mathbf{n}, \hat{\boldsymbol{\omega}}\left(\mathbf{y}\right), \tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\omega}}}\left(\mathbf{y}\right))$ from simulating toy experiments. How to perform this *sampling* of the test-statistic is dictated directly from the details of the sampling models for $\mathbf{n}$ and $\mathbf{y}$. One example is shown in Fig. 5.3 using the single powerlaw model (10 parameters, including detector model, cf. Sec 5.7) employed in this work, but with an exposure corresponding to 200 years of data taking. The sampling distribution of the un-penalized test-statistic (eq. (5.37)) corresponding to not including calibration data is shown in blue, while the penalized test-statistic distribution (eq. (5.40)) is shown in red. As expected from the discussions above, we observe a noticeable shift to larger values that in this simulated case correspond to 5 additional degrees of freedom[8] corresponding

[6] Note, eq. (5.40) is independent of any unknown parameters and is a function of observations only - as required for a valid test-statistic!

[7] $E[\cdot]$ denotes the average value over repeated measurements

[8] we simulated the real scenario of "sharing" 5 parameters with the calibration datasets, cf. Sec 5.8, especially Tab. 5.3, i.e. $\dim \boldsymbol{\theta} = 10$, $\dim \boldsymbol{\theta}_{-\boldsymbol{\omega}} = 5$

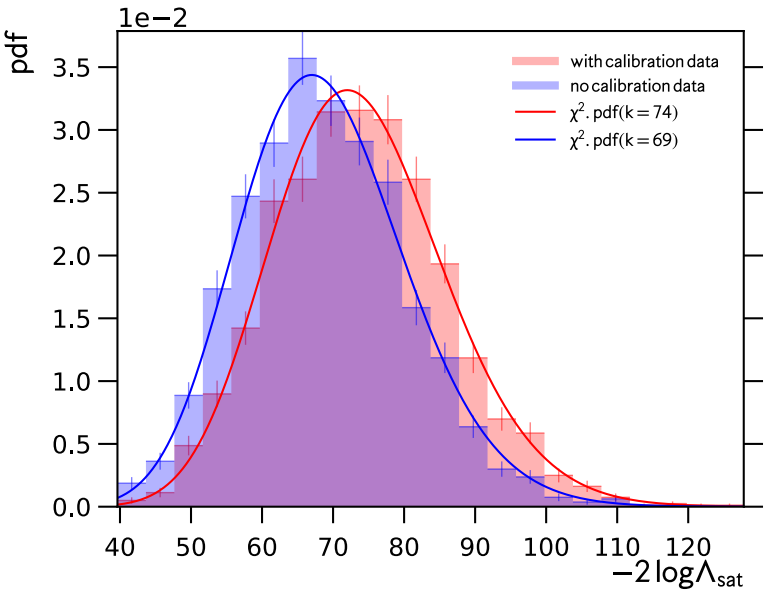


FIGURE 5.3: Distribution of the "saturated poisson likelihood" test-statistic $-2 \log \Lambda_{sat}$ used to loosely check for appropriate model fit: calculated from simulated toy experiments (histograms) and compared to chi-square distributions (solid lines). Distribution before inclusion of calibration data / constraints (blue) and after (red) - see text for details.

to the auxiliary datasets.

It is worth highlighting that eq. (5.40) remains a goodness-of-fit test for the modeling of the observations $\mathbf{n}$ (our sample) only, similar in spirit to eq. (5.37). This is because we are only relaxing the model M that connects $\mathbf{n}$ with the parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_{-\boldsymbol{\omega}}, \boldsymbol{\omega})$. The model M' that connects the shared nuisance parameters $\boldsymbol{\omega}$ to the auxiliary measurements $\mathbf{y}$ remains unquestioned. We are thus not testing the validity of the model assumptions that went into the calibrations measurements, as desired[9]. The construction of a full, joint goodness-of-fit test for $\tilde{M} = (M, M')$ on $\mathbf{z} = (\mathbf{n}, \mathbf{y})$ goes beyond the scope of (and is not of interest for) this thesis. Finally, a possible rejection of the model M due to large values of $\xi(\mathbf{n}, \hat{\boldsymbol{\omega}}(\mathbf{y}), \tilde{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\omega}}}(\mathbf{y}))$ can be due to two extremes. The data $\mathbf{n}$ is not well described by the "best fit" $\hat{\boldsymbol{\theta}}$ for model M, corresponding to large values of $\xi(\mathbf{n})$, or, assuming that M' is true, the best-fit values for the nuisance parameters $\hat{\boldsymbol{\omega}}(\mathbf{n}, \mathbf{y})$ are in conflict with the data $\mathbf{y}$. Of course combinations thereof are possible.

5.6 Binning of Observables

The binning of the relevant observables (energy, zenith angle, event type) adopted in this work is mostly based on considerations of systematic uncertainties and the expected

[9]The corresponding quality checks of the appropriateness of the modeling of the calibration data are obviously done elsewhere - in the context of the calibration measurements themselves

size of the data sample inferred from the 10%-burnsample. In general we prefer a large number of bins in energy over a large number of bins in zenith, because of the systematic uncertainties in the modeling of the directional reconstruction of cascades (cf. Sec. 4.1.2). In addition to energy an zenith bins, we chose three event type bins: *cascade sample*, *starting track sample*, and *single muon sample*, corresponding to the event selection discussed in Sec. 3.
The details of the binning are shown in Tab. 5.1 and are discussed below.

Single Muon Sample

The single muon sample contributes only a single bin to the likelihood function eq. (5.14). This choice is dictated by limited computing resources that prevent us from generating simulation datasets for the single muon background for varying assumptions of the detector response. We are therefore unable to appropriately model the effect of systematic uncertainties on the single muon contribution, in particular the shape of the its energy and zenith distributions. By limiting this sample to a single bin only we eliminate its sensitivity to shape changes due to detector systematics. This sample thus only constrains the effective overall normalization of the single muon component in the combined fit. The "bin" is defined as follows: $2.6 \leq \log_{10} E_{rec}/\mathrm{GeV} < 4.8$ and $-1.0 \leq \cos\theta_{rec} \leq 1.0$.

Starting Track Sample

The starting track sample contributes 12 bins to this analysis. They are defined in the range from $2.6 \leq \log_{10} E_{rec}/\mathrm{GeV} < 4.8$ with a width of $\Delta \log_{10} E_{rec}/\mathrm{GeV} = 0.2$ and $-1.0 \leq \cos\theta_{rec} \leq 1.0$, i.e. each bin contains events from the entire sky.

Cascade Sample

The cascade sample is the primary sample of this analysis. Here we use 66 bins in the range from $2.6 \leq \log_{10} E_{rec}/\mathrm{GeV} < 7.0$ with a width of $\Delta \log_{10} E_{rec}/\mathrm{GeV} = 0.2$, defined separately for up-going events (northern sky, $-1.0 \leq \cos\theta_{rec} < 0.2$), inclined down-going events (southern sky, $0.2 \leq \cos\theta_{rec} \leq 0.6$) and near vertically down-going events (southern sky, $0.6 < \cos\theta_{rec}$). The choice of using only three bins in the zenith angle is conservative. While the available data statistics would allow for more bins, we unfortunately discovered in this work that the reconstructed cascade directions are very sensitive to assumptions about the optical transport of photons in the ice. As we will discuss below, we are able to model the associated uncertainties using parameterizations obtained from simulations (see Sec. 5.7), but the available simulation statistics quickly becomes insufficient as the number of bins grows. This is especially true for near vertical directions.

Event Type Bin	NBins	Energy Range	Width	Zenith Range	Widths
cascade sample	66	2.6 – 7.0	0.2	−1.0 – 1.0	1.2, 0.4, 0.4
starting track sample	11	2.6 – 4.8	0.2	−1.0 – 1.0	2.0
single muon sample	1	2.6 – 4.8	2.2	−1.0 – 1.0	2.0

TABLE 5.1: The binning of observables (reconstructed energy and zenith) used in the maximum likelihood fit. Energy ranges are given in logarithmic units: $\log_{10} E/\mathrm{GeV}$

.

Systematic	Values
rel. DOM efficiency	0.81, 0.90, 0.95, **0.99**, 1.08
rel. scattering (bulk)	−10%, ±**0**%, +10%
rel. absorption (bulk)	−10%, ±**0**%, +10%
abs. scattering (HI)	30 cm, **50 cm**, 100 cm

TABLE 5.2: Simulated variations of assumptions about the detector response to light signals. The nominal values are marked in bold font.

5.7 Treatment of Systematic Uncertainties

The sources of systematic uncertainties and their impact on the observables of this analysis have been studied in Chapter 4. Here we discuss their statistical modeling and treatment in the maximum likelihood fit.

5.7.1 Detector Systematics

The effect of systematics variations of the detector model can only be evaluated at a few discrete points within the parameter space preferred by calibration measurements, due to the computationally expensive nature of the Monte Carlo simulations. The simulated systematics scenarios are shown in tab. 5.2. The technical details of how to include those uncertainties into a high-energy astrophysical neutrino flux measurement are the center of recurring discussions within the IceCube Collaboration. One possibility is to consider the systematics datasets as different *discrete detector models*, giving rise to *discrete nuisance parameters*. These methods are fully appropriate if the corresponding effect/process is truly discrete in nature. None of the detector systematics considered in this work fall into that category, however. They are intrinsically continuous, i.e. values in between the simulated scenarios are as plausible as the ones that were actually simulated. If the separation of the discrete points in the parameter space is large compared to the ability of the data to distinguish the two cases, such a treatment will reduce to the special case of not including systematic uncertainties at all. The extent to which this is appropriate depends on the problem at hand, e.g. even for a fixed dataset it will depend on what hypotheses and models of interest are being tested. Underestimating systematic uncertainties is a real possibility in such cases.
In order to avoid this issue we will model continuous effects with *continuous nuisance parameters*. This requires interpolating between the different discrete points at which the detector model was simulated. The likelihood function eq. (5.14) depends on the nuisance parameters only via the expected number of events μ_i in the observable bin i.

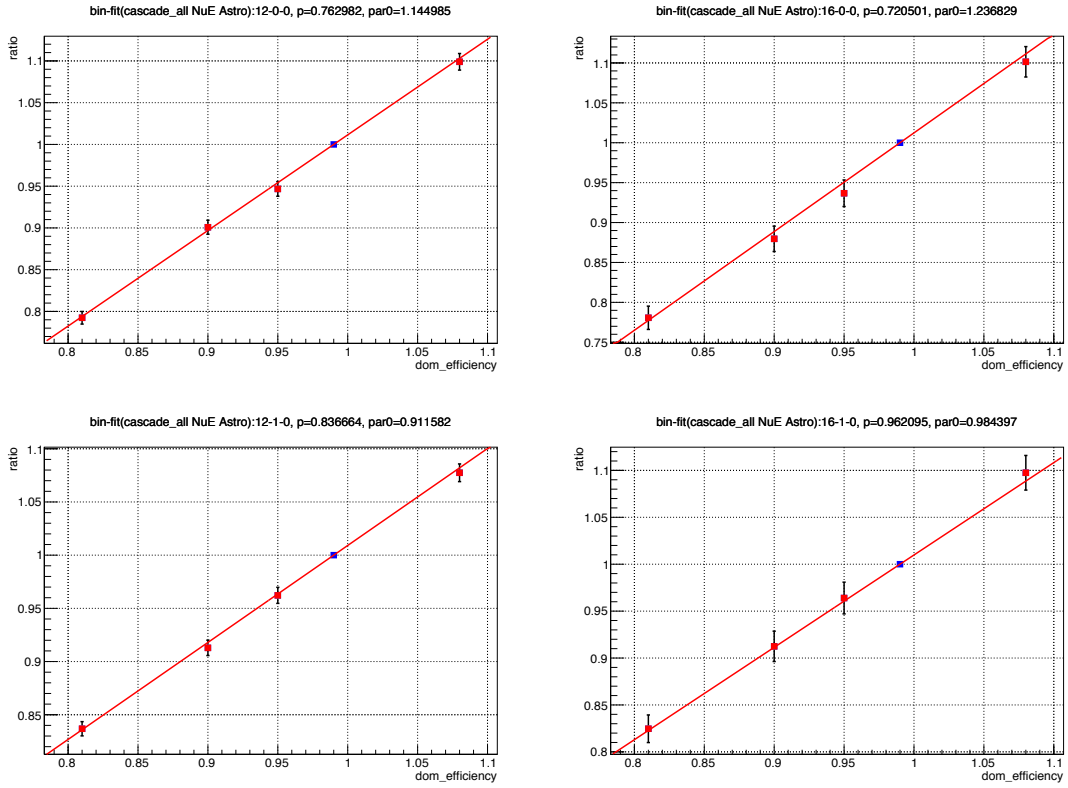


FIGURE 5.4: Per-bin parametrization of systematic effects on expected number of events relative to the nominal simulation in each bin. Here: impact of variations of assumed dom-efficiency on four example bins.

Interpolating between different discrete detector models thus becomes interpolating the relative changes in the expected number of events (efficiencies) $\epsilon_{i,j}(\omega_j)$ as function of the systematic nuisance parameter ω_j. The simplest relationship is linear:

$$\epsilon_{i,j}(\omega_j) = \mu_i(\omega_j)/\mu_i^0 \tag{5.42}$$

$$\epsilon_{i,j}(\omega_j) = 1.0 + a_{i,j} \times (\omega_j - \omega_j^0) \tag{5.43}$$

Assuming the efficiency corrections $\epsilon_{i,j}(\omega_j)$ for effect j to be independent from effect k for $(k \neq j)$, i.e. $\epsilon_{i,j}(\boldsymbol{\omega}) = \epsilon_{i,j}(\omega_j)$ we can write

$$\mu_i(\boldsymbol{\omega}) = \mu_i^0 \times \Pi_j \epsilon_{i,j}(\omega_j) \tag{5.44}$$

where μ_i^0 denotes the expected number of events in bin i when all systematics parameters are at their baseline values (nominal detector simulation). The coefficients $a_{i,j}$ can be obtained from fitting eq. (5.42) to the simulated datasets, individually for each bin i and nuisance parameter ω_j. The parametrization ensures that the model reduces the the nominal detector simulation for $\boldsymbol{\omega} = \boldsymbol{\omega}^0$, while it recovers the simulated systematics variation j with value $\omega_j = \omega_j^{sys}$ for $\boldsymbol{\omega} \equiv (\boldsymbol{\omega}_{-j}, \omega_j) = (\boldsymbol{\omega}^0_{-j}, \omega_j^{sys})$.

Fig. 5.4 shows the resulting parametrizations of the systematic DOM efficiency variations for four example bins. The simulated DOM efficiency variations yield linear efficiency

corrections within each bin, as expected. The other detector systematics are treated similarly. We verified our implementation by reproducing the discrete points at we which we simulated different detector configurations from our parametrizations obtained from the same simulations. The result is shown in Fig. 5.5 for electron neutrinos (top) and muon neutrinos (bottom) that contribute to the final cascade sample. We recover the intended models within deviations that are smaller than $\sim 1\%$ ($\sim 5\%$) for electron (muon) neutrinos. There are larger deficiencies at very small ($< 1\,\mathrm{TeV}$) or very high ($> 1\,\mathrm{PeV}$) energies. At these energies the bins either do not contribute to the likelihood ($\log_{10} E/\mathrm{GeV} < 2.6$) or the expected number of data events in such bins is sufficiently small that such departures are negligible compared to the statistical uncertainties in those bins.
Due to the large differences in the energy spectra of the different components (conventional atmospheric neutrinos, prompt atmospheric neutrinos and astrophysical neutrinos) and neutrino flavors it is necessary to perform the corresponding interpolations in each bin i separately for all of them.

$$\mu_i(\boldsymbol{\omega}) = \sum_{j\in\{e,\mu,\tau\}} \Big[\mu_{i,j}^{\mathrm{conv}\,\nu,0} \times \Pi_k \epsilon_{i,j,k}^{\mathrm{conv}\,\nu}(\omega_k) + \mu_{i,j}^{\mathrm{prompt}\,\nu,0} \times \Pi_k \epsilon_{i,j,k}^{\mathrm{prompt}\,\nu}(\omega_k) + \mu_{i,j}^{\mathrm{astro}\,\nu,0} \times \Pi_k \epsilon_{i,j,k}^{\mathrm{astro}\,\nu}(\omega_k)\Big] \tag{5.45}$$

Finally, in order to speed up the computations, we derive the parameterizations of the effects of the detector systematics on the astrophysical component in the energy spectrum only once, using the astrophysical best-fit parameters obtained from a fit based on the nominal detector modeling (no systematic variations). The parameterizations are then kept constant throughout the final fit including systematic uncertainties. Using the nominal astrophysical model, the single power law, we have verified that this has negligible impact on the results. Varying the astrophysical parameters assumed in the derivation of the parametrization within reason yields relative changes in the final maximum value of the log-likelihood function of $O(10^{-4})$, while relevant scales for the purpose of this measurement are $O(1)$. Observed differences in the best fit parameters are similarly negligible.

When parametrizing the impact of changes in the hole ice scattering lengths λ within the range given by our simulations (30 cm, 50 cm, 100 cm) we first transformed the scattering lengths into relative scattering coefficients, measured relative to the scale set by the default of 50 cm:

$$\alpha_{HI}(50\,\mathrm{cm}) \equiv 1.0 \tag{5.46}$$

$$\alpha_{HI}(\lambda) = \frac{50\,\mathrm{cm}}{\lambda} \tag{5.47}$$

This allows to treat hole-ice scattering on the same scale as bulk-ice scattering and absorption. On this relative scale the holeice scattering length of 30 cm corresponds to $\alpha_{HI} = 1.67$.

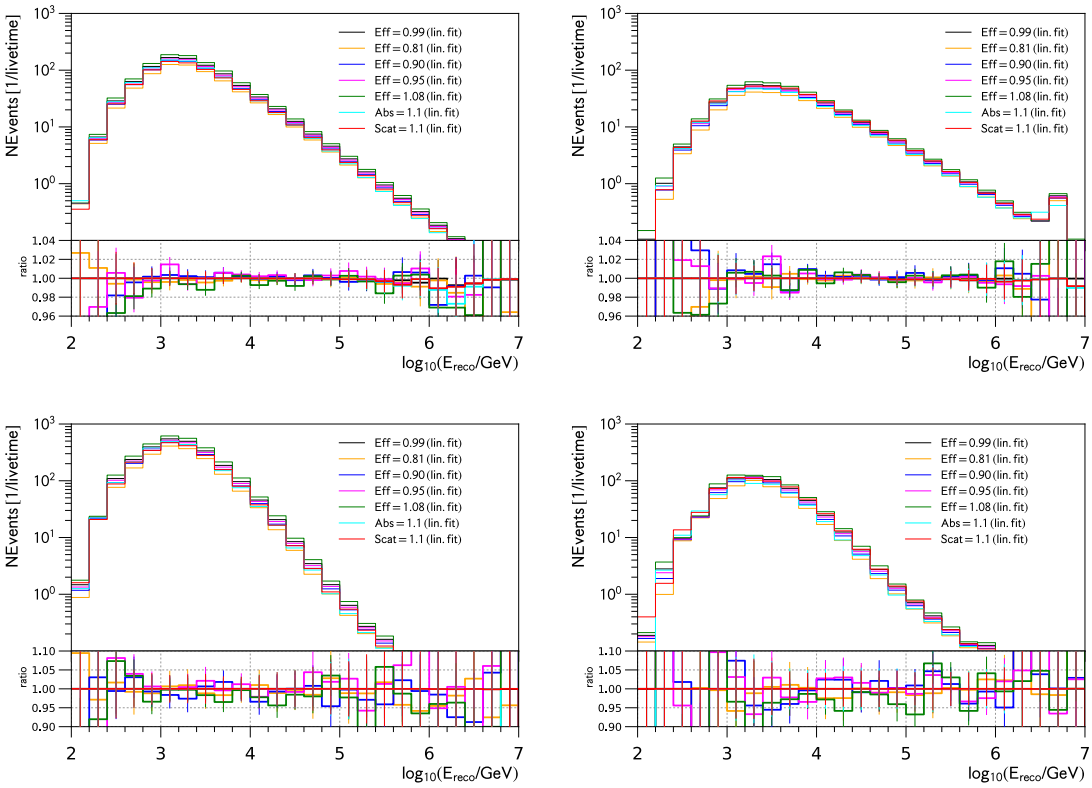


FIGURE 5.5: Comparison of (linearly) parametrized systematics to the simulation prediction used to create the parameterizations: electron neutrinos (top) and muon neutrinos (bottom). Up-going events (left) and down-going events (Right). The simulation prediction can be recovered within a $\sim 1\%$ ($\sim 3\%$) for electron neutrinos (muon neutrinos).

5.7.2 Atmospheric Background Systematics

Primary Cosmic Ray Spectrum

Uncertainties related to the primary cosmic-ray flux affect all three background components: atmospheric muons, atmospheric conventional as well as prompt neutrinos, as discussed in Sec. 4. Following previous IceCube measurements (e.g. [199]) we account for uncertainties in the energy spectrum of cosmic rays at energies relevant for this work by allowing for cosmic-ray spectral index γ_{CR} to vary within $\sigma_{\Delta\gamma_{CR}} = 0.05$. This is implemented using energy dependent correction factors to the event weights w_i^0 that are obtained from the nominal flux assumption for every event i:

$$w_i\left(E, \Delta\gamma_{CR}\right) = w_i^0 \times \left(\frac{E}{E_0}\right)^{-\Delta\gamma_{CR}} \tag{5.48}$$

E_0 denotes the median neutrino (muon) energy that is expected for this selection from the nominal flux model. For the three components we found:

- conventional ν: $E_0 = 1.8\,\text{TeV}$
- prompt ν: $E_0 = 5.1\,\text{TeV}$
- cosmic-ray μ: $E_0 = 2.2\,\text{TeV}$

Choosing E_0 as the median predicted neutrino (muon) energy, while fully general, reduces the correlation between $\Delta\gamma_{CR}$ and the overall normalization of the respective component in the final fit [200]. The impact on the observables has been discussed in Sec. 4.2.1.

Hadronic Interaction Models

At energies relevant for this work, the uncertainties in the conventional atmospheric neutrino flux due to uncertainties in the modeling of hadronic interaction is larger than that related to the primary cosmic ray flux (c.f. Sec. 4). Unfortunately there does not appear to be an "obvious" prescription on how they best be treated. In Sec. 5.7.1 we argued that continuous modeling of continuous systematic uncertainties should be preferred over a discrete statistical treatment to avoid underestimating them. In the context of hadronic interaction models it is not obvious how such a continuous model could be achieved. As we will see later on, hadronic interaction uncertainties are subdominant to the detector related ones. We therefore decided to treat them as discrete possibilities. When construction confidence intervals (c.f. Sec. 5.4) and contours for our fit parameters, we will repeat the calculation for the different hadronic interaction models. The combined likelihood will at each point in parameter space be given by the minimum likelihood value from the set of values obtained with the discrete choices. The following post-LHC hadronic interaction models will be used:

- DPMJETIII v.17.1
- QGSJETII v.04
- EPOS LHC v.XX
- SIBYLL v.2.3c
- HKKMS06[10]

Prompt Atmospheric Neutrinos

We account for systematic uncertainties in the theoretical calculation of the prompt atmospheric neutrino flux by allowing its total normalization to be a free parameter in the fit. Since the contribution from prompt neutrinos to this sample is expected to be subdominant at all energies and zenith angles, small differences in the spectral shape of the prompt neutrino flux can be ignored.

[10] The hadronic interaction model used by HKKMS06 in [99] is a modification based on DPMJETIII.

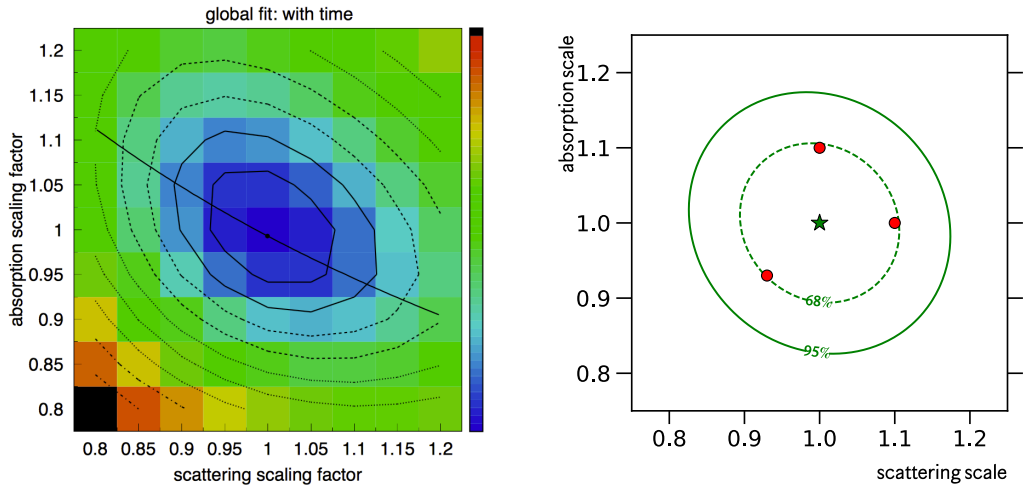


FIGURE 5.6: Constraints on global changes in scattering and absorption of the ice (from calibration data [125] (left). Implementation as bi-variate normal measurement error into this analysis (right). Red points mark available systematics simulation datasets.

5.8 Parameter Constraints from Calibration Data and Prior Information

As discussed in Sec. 5.3 we model the parameter constraints obtained from independent calibration measurements using a multi-variate normal[11] measurement model. In this particular case they can be sub-divided into three independent univariate normal distributions (dom efficiency, hole ice scattering scale and cosmic-ray spectral index deviation) and one bivariate normal distribution (optical properties of bulk ice: scattering and absorption). The corresponding contribution to the likelihood function eq. (5.20) reads

$$\begin{aligned}\frac{1}{2}\left(\hat{\boldsymbol{\omega}}-\boldsymbol{\omega}\right)^{T}\boldsymbol{\Sigma}^{-1}\left(\hat{\boldsymbol{\omega}}-\boldsymbol{\omega}\right) = \frac{1}{2}\left(\frac{\epsilon_{doms}-\hat{\epsilon}_{doms}}{\sigma_{\epsilon}}\right)^{2} + \frac{1}{2}\left(\frac{\alpha_{HI}-\hat{\alpha}_{HI}}{\sigma_{\alpha}}\right)^{2} \\ + \frac{1}{2}\left(\frac{\Delta\gamma-\hat{\Delta}\gamma}{\sigma_{\Delta\gamma}}\right)^{2} + \frac{1}{2}\left(\boldsymbol{\alpha_{BI}}-\hat{\boldsymbol{\alpha}}_{\boldsymbol{BI}}\right)^{T}\boldsymbol{\Sigma}_{BI}^{-1}\left(\boldsymbol{\alpha_{BI}}-\hat{\boldsymbol{\alpha}}_{\boldsymbol{BI}}\right)\end{aligned} \tag{5.49}$$

The numerical values for the various quantities $\hat{\xi}$ and σ_{ξ} are given in Tab. 5.3 and follow from the discussions in Sec. 4 and Sec. 5.7. The bivariate treatment of the optical properties of bulk-ice is needed, since the measurement of the absorption and scattering scales revealed anti-correlation between those two parameters. Fig. 5.6 shows the original measurement (left) and our bivariate approximation (right) based on input of IceCube's calibration group (three red points)[12]. Our approximation is characterized by a correlation coefficient of $\rho = -0.1$ and marginal standard deviations of $\sigma_i = 0.07$.

[11] In practice we truncated the allowed values of some of the nuisance parameters at zero.
[12] Three points are sufficient to uniquely define an ellipse of its center is known.

Nuisance Parameter	Symbol	Calibration Result	Calibration Uncertainty
rel. dom efficiency	ϵ_{dom}	0.99	0.10
rel. scattering (hole ice)	α_{HI}	1.00	0.67
rel. absorption (bulk ice)	α_{abs}	1.00	0.07 (*)
rel. scattering (bulk ice)	α_{scat}	1.00	0.07 (*)
cosmic-ray index deviation	$\Delta\gamma_{CR}$	0.0	0.05
muon flux norm	Φ_{μ}	–	–
conv flux norm	Φ_{conv}	–	–
prompt flux norm	Φ_{prompt}	–	–

TABLE 5.3: Summary of constraints (uncertainties) on systematics related nuisance parameters from calibration measurements.

IceCube has attempted to measure the prompt neutrino flux several times in the past [164][201]. Since no prompt flux was observed, limits on its normalization were obtained. In particular measurements using through-going muon neutrinos from the northern sky derived strong constraints. It is very tempting to include these previous studies as independent calibration measurements. This is problematic. As discussed in Sec. 5.3 the corresponding penalty terms are simply an approximation to a joint analysis of both datasets (main dataset + calibration dataset). If the two studies share additional parameters they would thus need to be included as well. In other words eq. (5.20) only applies if all shared parameters required to model datasets $\mathbf{x}$ and $\mathbf{y}$ are accounted for, so they can be jointly maximized during subsequent analysis. In particular eq. (5.20) does not apply if a subset of the shared parameters has been "removed" from the calibration result via explicit profiling - which is the case if only the prompt limit was used. One could attempt to construct a multi-variate parametrization of the likelihood function that characterizes the information contained in the previous measurement. Because of the large number of parameters involved, we believe this to be hopeless and that it better be left to an explicit joint analysis in form of a global fit (e.g. [199][202]). Additional complications arise when both measurements disagree on some of the shared parameters that could be indicative of insufficiencies in the model itself. For example the astrophysical neutrino spectra inferred from measurements using through-going muons from the northern sky appear in mild tension with those obtained from cascade dominated measurements that are more sensitive to the Southern Sky at lower neutrino energies. We thus decided not to include any constraints on the prompt neutrino flux normalization that are rooted in previous IceCube measurements.

Chapter 6

Results

According to our analysis the data sample is well described by the statistical model introduced in Chapter 5 and corresponding assumptions, including the choice of modeling the astrophysical neutrino flux using a single, unbroken powerlaw. In particular we find

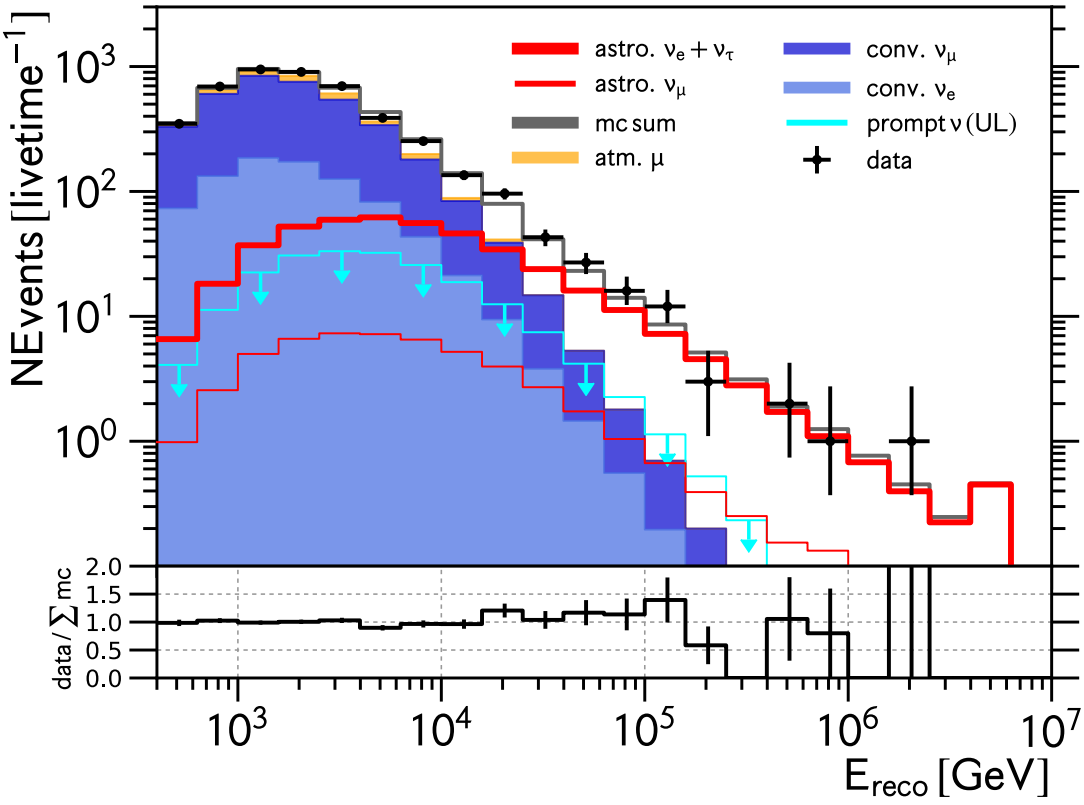


FIGURE 6.1: Best-Fit reconstructed cascade energy spectrum: estimated contribution from astrophysical electron and tau neutrinos (thick red line) as well as muon neutrinos (thin red line). Upper-limit on prompt atmospheric neutrino contribution (light blue) at 90% C.L.

a soft spectral index

$$\gamma_{astro} = 2.53^{+0.07}_{-0.09} \tag{6.1}$$

and a total astrophysical neutrino flux per neutrino flavor

$$\Phi^{100\,\mathrm{TeV}}_{\nu_x+\bar{\nu}_x} = (1.58^{+0.25}_{-0.28}) \times 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1} \tag{6.2}$$

at a neutrino energy of 100 TeV. The best-fit values and uncertainties obtained for all parameters are given in Tab. 6.1. Fig. 6.1 shows the corresponding energy spectrum for all events in the cascade sample, i.e. events consistent with cascade-like light emission and arrival directions from the entire sky. As described in Sec. 5.6, the fit is performed after grouping events additionally according to arrival direction into three bins. The corresponding cascade energy spectra for near vertically down-going ($\cos\theta_{rec} > 0.6$), near horizontally down-going ($0.2 < \cos\theta_{rec} \leq 0.6$) and all up-going trajectories ($\cos\theta_{rec} \leq 0.2$) are shown in Fig. 6.2 and Fig. 6.3 (top left and top right) respectively. The spectrum obtained with the control sample of starting tracks is shown in Fig. 6.3 (2nd row, left) as well. We observe excellent agreement between the data the expectations from the best-fit as evidenced by a corresponding goodness-of-fit p-value of 0.59 (c.f. Sec. 5.5), obtained from repeatedly performing the fit on replicated, artificial datasets, called toy-experiments hereafter [1], as shown in Fig. 6.3 (2nd row, right). The agreement is observed in all four spectra, even the nearly vertically down-going selection, despite very different composition (astrophysical signal, atmospheric backgrounds, event signatures) and therefore different sensitivity to the various systematic uncertainties identified to be relevant for this work.

All spectra show the contribution from astrophysical electron and tau neutrinos (cascades) as thick red line, while the contribution from astrophysical muon neutrinos (tracks) is marked by the thin red line. As expected by design of this analysis, the contribution

[1] This procedure is known as *parametric bootstrap* in statistics.

Parameter		Calibration	Result $\pm 1\sigma$ ($< 90\%$ upper limit)
spectral index	γ	-	$\mathbf{2.53^{+0.07}_{-0.08}}$
norm astro	ϕ	-	$(\mathbf{1.58 \pm 0.25})$ c.u.
norm conv	ϕ_{conv}	-	$(1.02^{+0.13}_{-0.11}) \times \Phi_{HKKMS06}$
norm prompt	ϕ_{prompt}	-	$< 4.8 \times \Phi_{BERSS}$
norm muon	ϕ_{muon}	-	1.42 ± 0.04
cosmic ray index dev.	$\Delta\gamma_{CR}$	0.00 ± 0.05	0.04 ± 0.03
scattering scale (bulk)	ϵ_{scat}	1.00 ± 0.07(*)	1.03 ± 0.03
absorption scale (bulk)	ϵ_{abs}	1.00 ± 0.07(*)	1.00 ± 0.04
scattering scale (HI)	ϵ_{abs}	1.00 ± 0.67	$1.68^{+0.19}_{-0.18}$
DOM efficiency	ϵ_{eff}	0.99 ± 0.10	$1.04^{+0.08}_{-0.07}$
$-2\log\Lambda_{sat}$	65.03	**p-value**	**0.59**

TABLE 6.1: Best-Fit values for all parameters, assuming the single power-law astrophysical flux model and the default conventional flux model (HKKMS06 [99]). Therefore the uncertainties on ϕ, γ do not yet include hadronic interaction uncertainties. (*) calibration constraints are anti-correlated, c.f. Fig. 5.6.

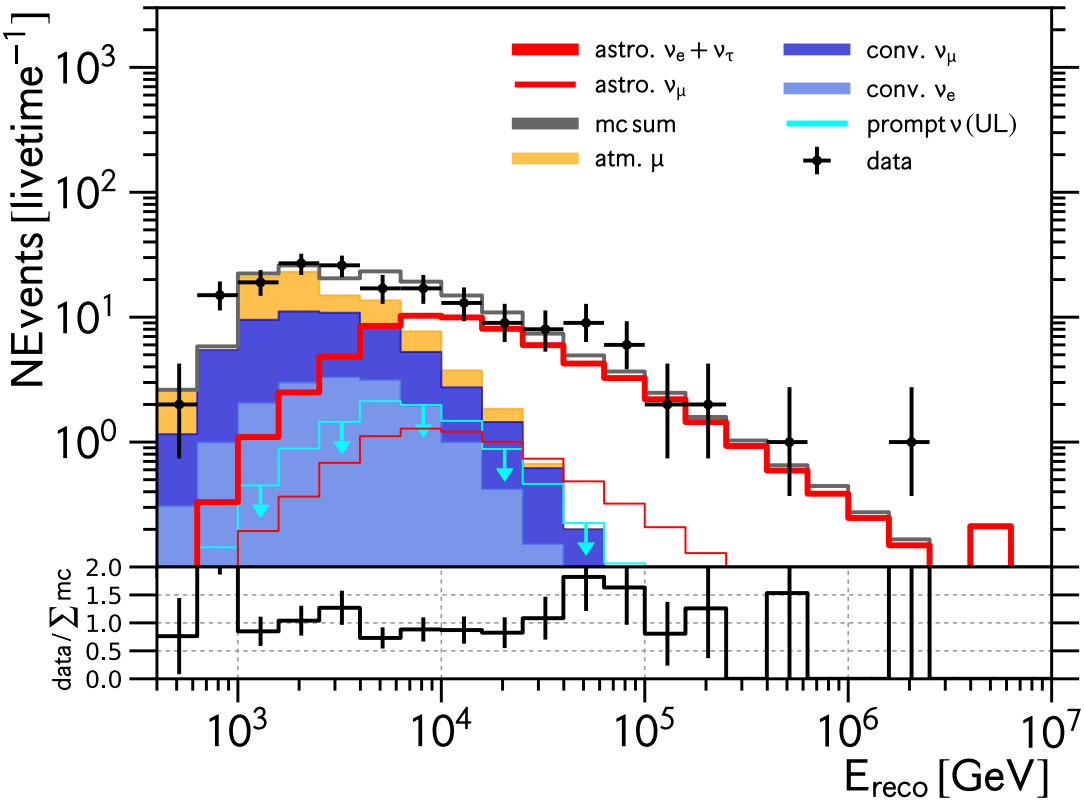


FIGURE 6.2: Best-Fit reconstructed cascade energy spectrum ($\cos\theta_{rec} > 0.6$): estimated contribution from astrophysical electron and tau neutrinos (thick red line) as well as muon neutrinos (thin red line). Upper-limit on prompt atmospheric neutrino contribution (light blue) at 90% C.L.

from muon neutrinos to the cascade sample, containing almost all of the astrophysical neutrinos, is strongly dominated by the contribution from electron and tau neutrinos. Correspondingly, the spectral properties of the astrophysical neutrino flux, obtained here, apply predominantly to the astrophysical electron and tau neutrino flux. The ability to distinguish event-types (cascades, starting tracks) is most visible in the starting track spectrum, Fig. 6.3 (2nd row, left), which is mostly composed of conventional atmospheric muon neutrinos (dark blue) with negligible contribution from conventional atmospheric electron neutrinos (light blue). While subdominant overall and in stark contrast to the cascade samples, the starting track spectrum is estimated to contain more astrophysical muon neutrinos than electron and tau neutrinos. According to our best-fit, when averaged across the entire sky, astrophysical electron and tau neutrinos dominate the observed cascade energy spectrum above reconstructed energies of $\sim 25\,\text{TeV}$ over atmospheric backgrounds (atmospheric muons, conventional and prompt atmospheric neutrinos). They remain dominant at even lower energies, if only near-vertically down-going events are considered. The atmospheric self-veto effect (c.f. Sec. 4.2.3) significantly reduces the atmospheric neutrino background and astrophysical neutrinos are estimated to remain dominant down to reconstructed energies of $\sim 6\,\text{TeV}$ ($\cos\theta > 0.6$). Fig. 6.4 (top) shows the zenith distribution of all cascade events with reconstructed energies above 10 TeV.

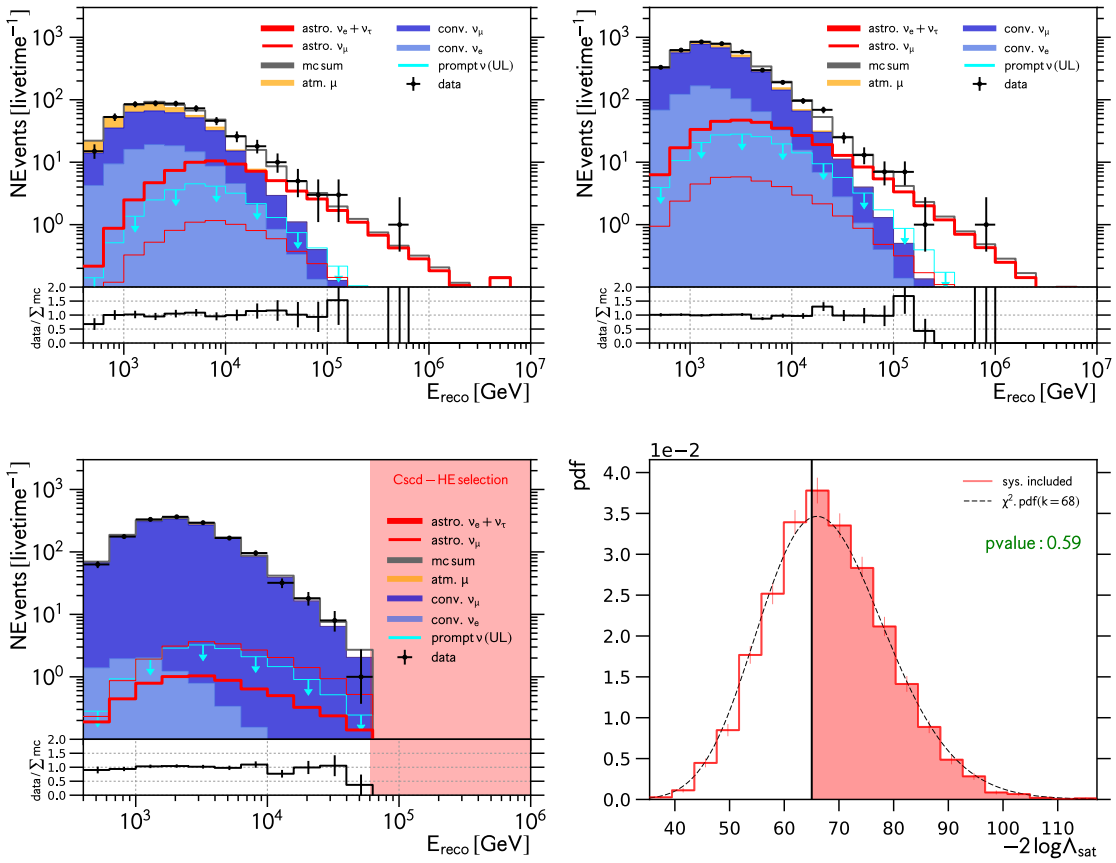


FIGURE 6.3: Best-Fit reconstructed cascade energy spectrum: $0.2 < \cos\theta_{rec} \leq 0.6$ (top left), $-1.0 \leq \cos\theta_{rec} \leq 0.2$ (top right). Best Fit reconstructed starting track energy spectrum (bottom left): estimated contribution from astrophysical electron and tau neutrinos (thick red line) as well as muon neutrinos (thin red line). Upper-limit on prompt atmospheric neutrino contribution (light blue) at 90% C.L. Goodness-of-Fit test-statistic distribution from toy experiments (bottom right)

The zenith dependence of the atmospheric self-veto effect manifests itself as a zenith dependence in the ratio between astrophysical (red) and atmospheric neutrinos (blue). It is worth pointing out that the shape agreement, especially in the up-going region, between data (black) and simulation (grey) can to some degree be regarded as independent confirmation of our best-fit model, since the fit itself was performed in three zenith bins only. The overall agreement remains good if the energy threshold is varied, as shown in Fig. 6.4 (middle left: $E > 60\,\mathrm{TeV}$, middle right: $E > 25\,\mathrm{TeV}$, bottom left: $E > 5\,\mathrm{TeV}$, bottom right: all energies). However at low energies (bottom row) and for very vertically up-going trajectories ($-1.0 \leq cos\theta_{rec} \leq -0.8$) we observe a significant departure of the number of observed events from the best-fit expectation for as of yet unknown reasons, potentially related to the refrozen ice in IceCube's drill holes. Since the corresponding contribution from that region to the up-going cascade sample ($cos\theta_{rec} < 0.2$) is small ($\sim 18\%$), this model deficiency does not appear to impact the overall goodness-of-fit of the result and thus this measurement of astrophysical neutrinos as such. We suspect that the in-situ PMT acceptance for vertical photon trajectories (i.e. head-on illumination) might not be

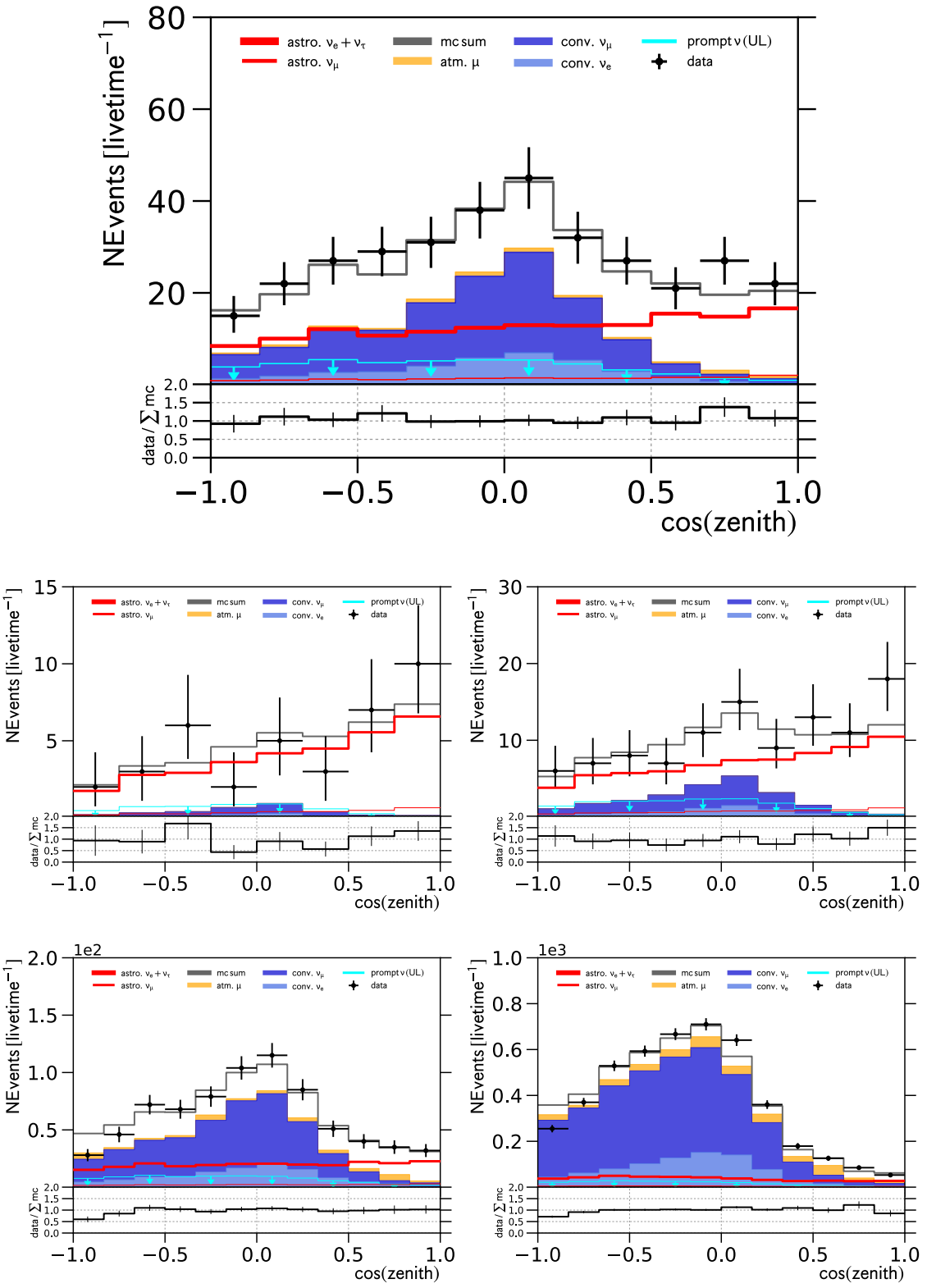


FIGURE 6.4: Best-Fit reconstructed cascade zenith distributions: $E_{rec} > 10\,\text{TeV}$ (top), $E_{rec} > 60\,\text{TeV}$ (middle left), $E_{rec} > 25\,\text{TeV}$ (middle right), $E_{rec} > 5\,\text{TeV}$ (bottom left), all energies (bottom right).

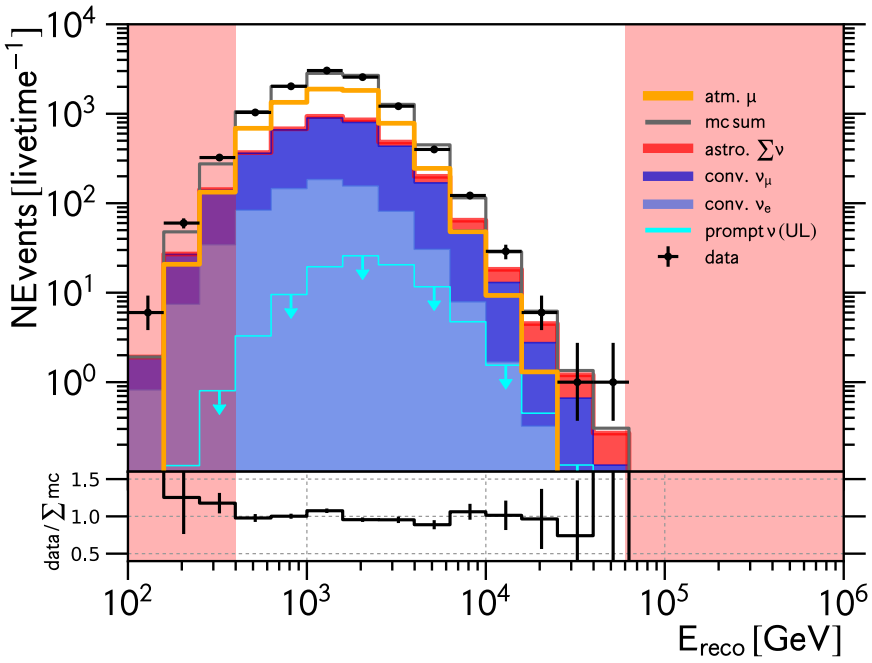


FIGURE 6.5: Best-Fit reconstructed energy spectrum in single muon sample: estimated contribution from astrophysical electron and tau neutrinos (thick red line) as well as muon neutrinos (thin red line). Upper-limit on prompt atmospheric neutrino contribution (light blue) at 90% C.L.. Only the normalization is used in the fit.

modeled sufficiently accurate. In fact, as pointed out in [203], such photon trajectories can not be probed well though IceCube's standard LED calibration system. Simulations with varying photon efficiencies as function of photon impact angle are currently being performed to test this.

The contribution from atmospheric muon background to the two neutrino samples, cascades and starting tracks, has been estimated to $\sim$ 8% and $\sim$ 3%, respectively. The contribution from atmospheric muons is only noticeable in the cascade sample for near-vertical trajectories at lowest energies, as shown in Fig. 6.2. The appropriateness of the modeling of the single muon energy spectrum can be checked much better using the muon control sample. In this sample the muon contribution is estimated to be $\sim$ 65% and, given the large statistics, allows for a clearer picture. The corresponding energy spectrum is shown in Fig. 6.5 and good agreement between data and simulation is observed. Note that in the fit we only use the total number of events (i.e. no shape information), since, due to the lack of sufficient simulation, we were unable to parametrize the corresponding systematic uncertainties (c.f. Sec. 5.6). The observed level of shape agreement can thus be viewed as independent confirmation of the method.

The main result of the statistical analysis assuming the astrophysical flux to follow a single, unbroken powerlaw is shown in Fig. 6.6 (top left): the two-dimensional confidence region of the two parameters Φ_{astro} and γ_{astro} (blue) in comparison to the expectation from the Asimov dataset (red) and an alternative fit performed in only two zenith bins separated at $\cos\theta = 0$ (green). All three contours appear very similar, which means

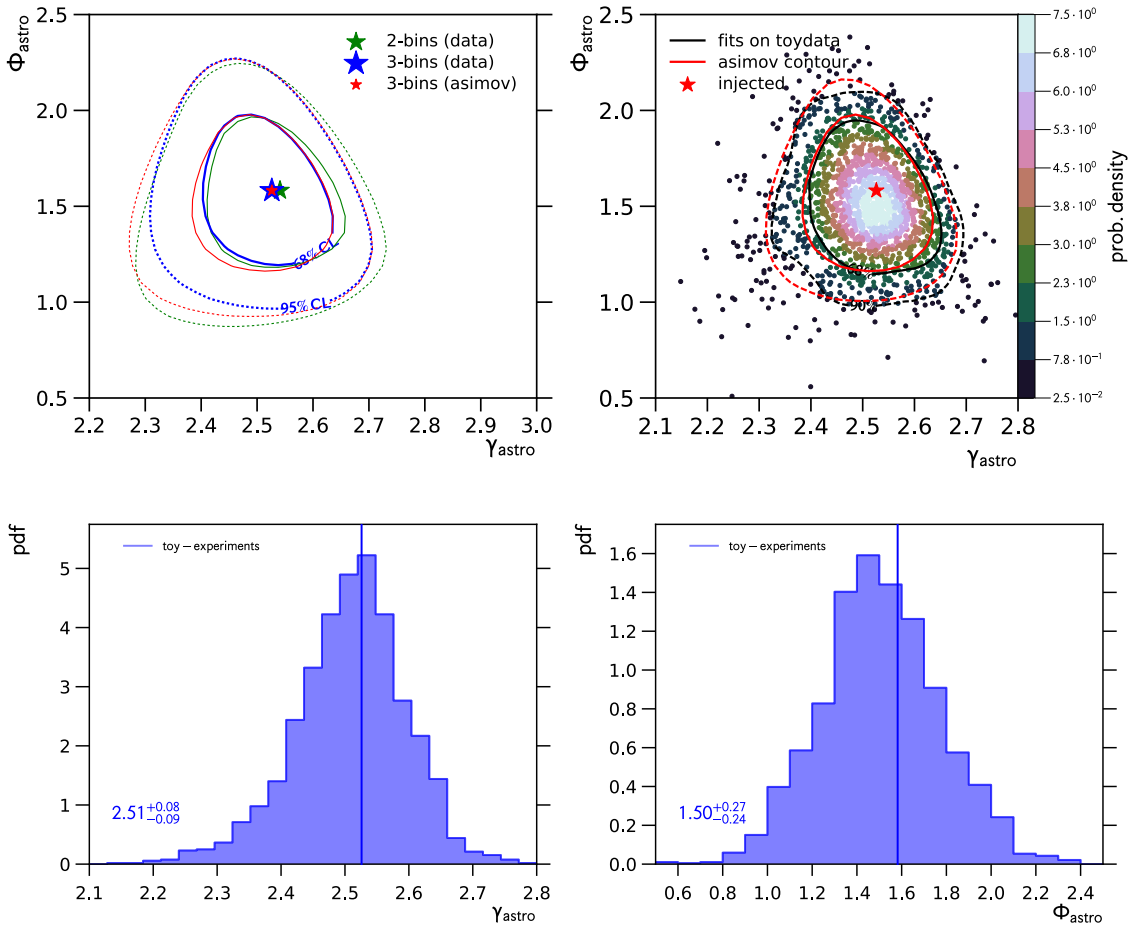


FIGURE 6.6: Top: Confidence contours for astrophysical flux parameters (left) observed in the data (blue) and median expectation (red). Two-dimensional sampling distribution of the best-fit values in comparison to median expected confidence contour (right). Bottom: Sampling distributions of astrophysical parameters: γ (left) and Φ (right).

that this measurement delivered the expected precision and the result appears stable against reasonable changes in the binning of the reconstructed zenith direction. Since the Asimov dataset only provides a statement about expected (median) test-statistic values [192], we studied the sampling distribution of the astrophysical best-fit values, assuming that the true values realized in nature coincide with the ones we have actually measured. This is again done by repeating the entire statistical analysis on replicated toy experiments. The result is shown in Fig. 6.6 (top right), where each point represents the solution obtained from one particular replication. The previously mentioned expected confidence contours from the Asimov dataset are shown in red. In addition we estimate equal-density contours that contain 68% (90%) of the replicated measurement results. They are shown in black and coincide reasonable well with the ones inferred from the Asimov dataset. We do, however, notice that the distribution of the replicated results appears slightly shifted to lower astrophysical normalization compared to the injected true value (red star). This is better visualized in Fig. 6.6 (bottom right). On average we

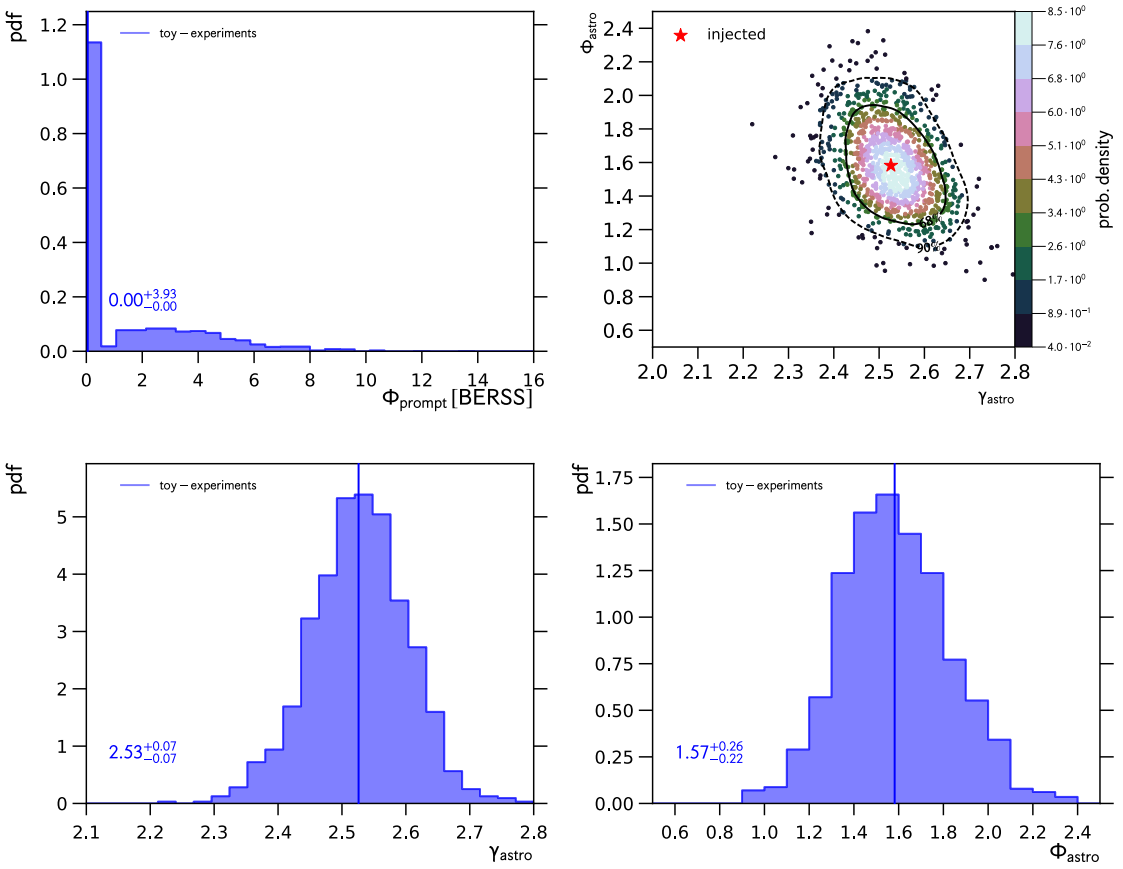


FIGURE 6.7: Sampling distribution of prompt atmospheric neutrino flux normalization Φ_{prompt} (top left). Two-dimensional conditional sampling distribution of astrophysical flux parameters ($\hat{\Phi}_{prompt} > 0$) (top right). Corresponding one-dimensional conditional sampling distributions (bottom).

would expect to measure a astrophysical normalization of $1.50 \cdot 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$ compared to the true value of $1.58 \cdot 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$ (vertical blue line, usual units), which implies the presence of a small (relative to the total variance) bias of -5% in the normalization measurement - if the true values realized in nature were given by our actual best fit (see Tab. 6.1). The astrophysical spectral index measurement instead appears approximately unbiased, as shown in 6.6 (bottom left). From a mathematical perspective, the existence of such a bias is not necessarily surprising, since maximum likelihood estimators do not guarantee unbiasedness for finite samples. This sample, however, is sufficiently large that several large sample theorems hold true, for example Wilk's theorem (c.f. Sec. 5.4) and, furthermore, such a bias has not been discussed in the context of previous IceCube measurements. The observed bias can be understood by focusing on the sampling behavior of best-fit values for the Φ_{prompt} parameter, the normalization of the prompt atmospheric neutrino contribution for which we injected $\Phi_{prompt} = 0 \cdot \Phi_{BERSS}$ as true value. The distribution is shown in Fig. 6.7 (top left) and consists of two components: a sharp peak at/near $\Phi_{prompt} = 0 \cdot \Phi_{BERSS}$ ($\sim 60\%$ of the results) and a large tail extending up to $\Phi_{prompt} = 10 \cdot \Phi_{BERSS}$. If we only consider the solutions that contribute to the peak at zero, the bias in the corresponding results for

the astrophysical normalization vanishes, see Fig. 6.7 (top right, bottom row) and the correct astrophysical normalization is recovered. Thus the bias is a boundary effect. We restricted the parameter space of the prompt normalization to positive values and hence only upward fluctuations can produce non-zero best-fit values leading to a reduction in the measured astrophysical normalization in order to preserve the total number of events. Due to the boundary restriction, this effect can not be balanced by downward fluctuations in the data. Thus, the bias is understood and, in particular, does not indicate an issue with our measurement. We will come back to similar non-trivial boundary effects in the context of hypothesis tests for additional complexity in the astrophysical flux model (c.f. Sec. 6.6). Finally, we would like to point out that the inferred sample standard deviations of $\sigma_\gamma = 0.07$ and $\sigma_\Phi = 0.24$ compare well to the uncertainties that we determined for the actually observed dataset (see Tab. 6.1).

6.1 Impact of Systematic Uncertainties

The relevant sources of systematic uncertainties and their implementation in this statistical analysis are discussed in Sec. 4 and 5.7. Fig. 6.8 (top) shows the correlation matrix visualizing the correlations among all 10 parameters in the fit. Since this matrix is symmetric, only the main diagonal and the lower triangle are shown. Of particular interest are correlations between the parameters of interest (Φ_{astro}, γ_{astro}) and the nuisance parameters that encode systematic uncertainties. We find the astrophysical spectral index γ_{astro} to be only weakly correlated with most parameters, except for the prompt atmospheric and astrophysical neutrino flux normalizations. Compared to the spectral index, the astrophysical flux normalization Φ_{astro} shows stronger (anti)correlations with various nuisance parameters, most prominently the conventional atmospheric neutrino normalization and the relative photon detection efficiency of the IceCube optical modules. We studied the relationships between the parameters of interest and the nuisance parameters in more detail by first varying the parameter of interest and then re-fitting the nuisance parameters. Subsequently we repeat the same for the nuisance parameters. The result is shown in Fig. 6.8 (middle and bottom rows). Increasing the value for the astrophysical normalization also increases the inferred conventional normalization but decreases the required relative efficiency of the DOMs to photons. One might wonder how increasing the background normalization can simultaneously increase the signal normalization. This is possible because signal and background contributions dominate at different energies (see Fig. 6.1) and thus are not expected to strongly influence each other. The astrophysical normalization is mainly influenced by variations in the relative DOM efficiency, because it is defined at a neutrino energy of $E_\nu = 100\,\text{TeV}$ but the mapping between true neutrino energy and reconstructed neutrino energy depends on the true value of the DOM efficiency (c.f. Sec. 4). Altering the DOM efficiency, however, also has a strong influence on the selection threshold of $Q_{tot} = 100\,\text{p.e}$ and therefore influences the normalization of the conventional background at low energies. This explains the strong anti-correlation ($\rho = -0.9$) that is observed between conventional normalization and DOM efficiency (see

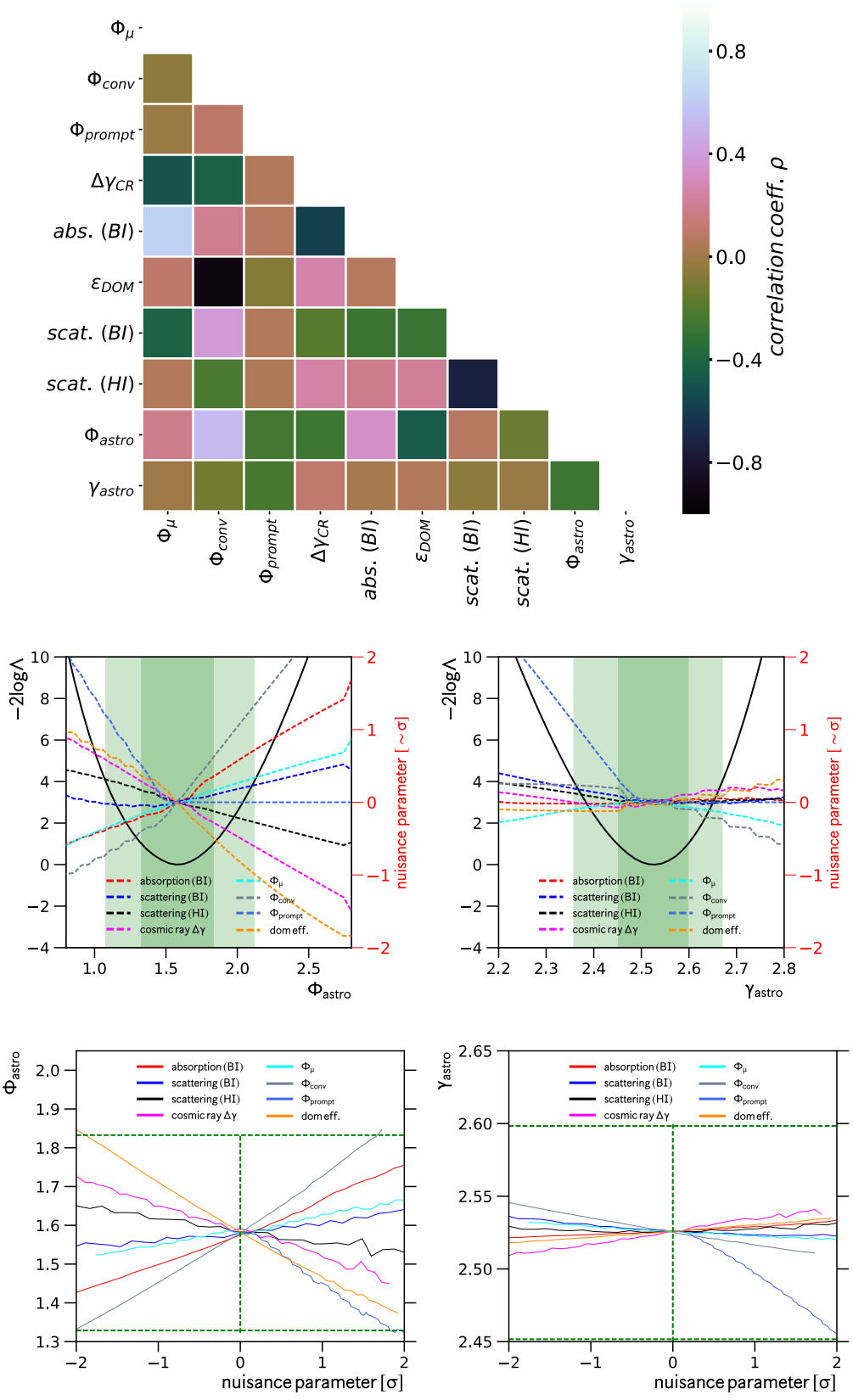


FIGURE 6.8: Correlation matrix of all fit parameters (top). Correlations between parameters of interest Φ_{astro}, γ_{astro} with nuisance parameters (middle, bottom).

Fig. 6.8, top) and demonstrates the importance of interactions across multiple fit parameters[2].
When the astrophysical normalization is decreased the prompt astrophysical normalization becomes important, since, at least for arrival directions that are not impacted by self-veto suppression, it adds to the total number of expected events at high energies, where conventional atmospheric neutrinos do not contribute. The influence of the prompt atmospheric neutrino flux is more visible for the astrophysical spectral index γ (see Fig. 6.8, bottom right). Increasing its value from the best fit of $\Phi_{prompt} = 0 \cdot \Phi_{BERSS}$ by 2σ would harden (decrease) the astrophysical spectral index to $\gamma = 2.45$.

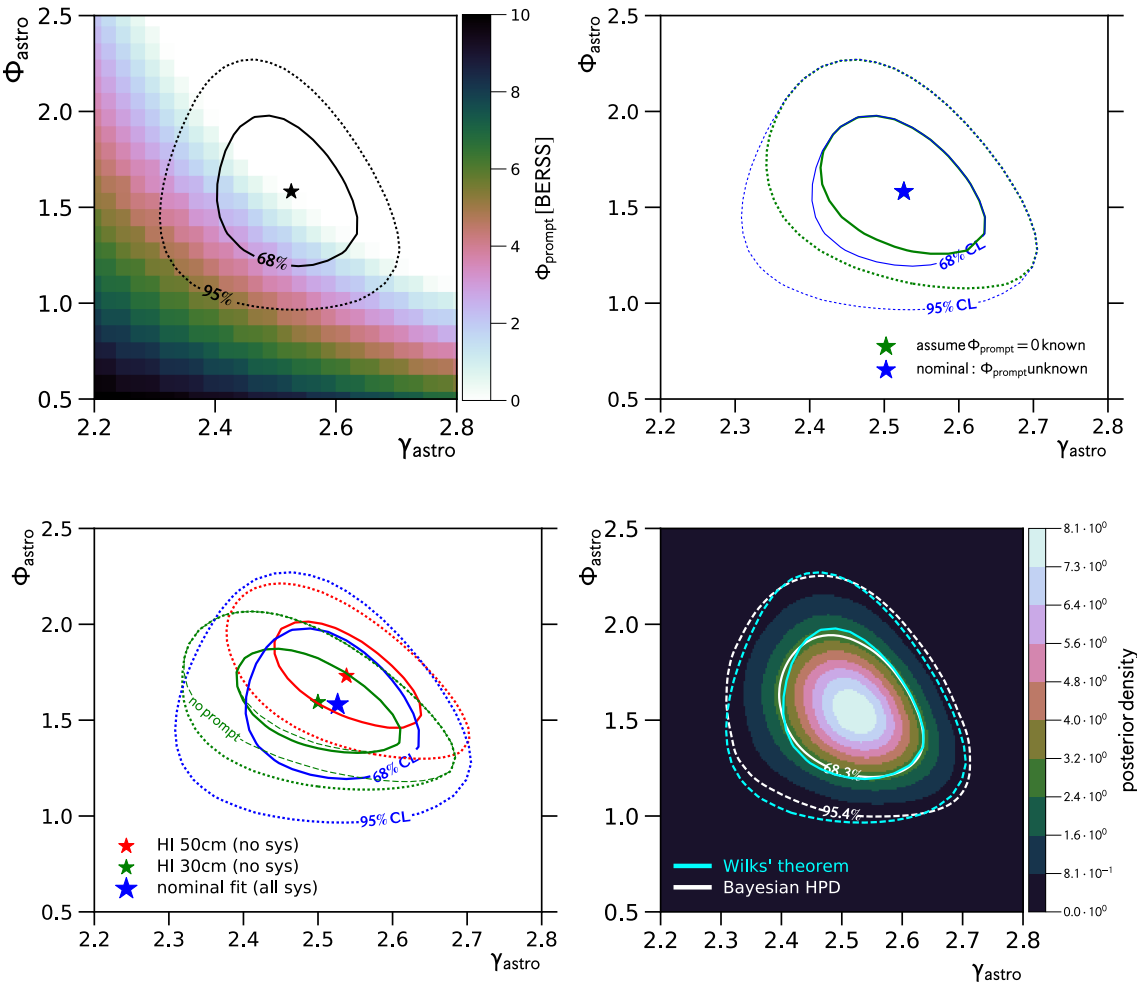


FIGURE 6.9: Impact of systematic uncertainties on confidence contour on astrophysical parameters of interest. Top: impact of prompt normalization. Bottom: impact of hole-ice and other detector systematics (left); Comparison of Wilks' contours to Bayesian HPD contours (right).

[2]In fact, when designing this analysis one could have easily argued, given their large anti-correlation, not to include both parameters, Φ_{conv} and ϵ_{DOM}, and instead use a single "effective" parameter. From the discussion it should be clear that his would have failed to account for an important contribution to the total systematic uncertainty in the measurement of Φ_{astro}.

Fig. 6.9 (top left) shows the preferred value of the atmospheric prompt neutrino flux normalization as a function of both astrophysical parameters. As expected from the discussions above, towards astrophysical fluxes with less intensity and harder spectral shape, the data requires the presence of a soft atmospheric prompt flux with increasing normalization. Fig. 6.9 (top right) compares the nominal two-dimensional confidence region for the astrophysical parameters (blue) to an alternative one, that assumes an (unrealistic) universe without prompt atmospheric component (green). The latter is smaller for obvious reasons and the difference between both, while not large compared to the total area, can be regarded as the contribution from systematic uncertainty related to the strength of the prompt atmospheric flux to the total uncertainty in the astrophysical flux measurement. The impact of the remaining nuisance parameters on the contour is shown in Fig. 6.9 (bottom left). Here, the green contour corresponds to a scenario in which all detector-related systematics parameters are known and kept fixed at their nominal values - except for the scattering length of the hole-ice which we fixed to 30 cm. Thus, only statistical uncertainties and uncertainties related to the intensities of the background fluxes contribute. The contour is noticeably smaller than the nominal one (blue). Also shown, in red, is the same calculation but assuming a detector with a 50 cm scattering length inside the bore holes. While strongly disfavored by the available data (cf. Sec. A.2) we observe a relatively small ($< 1\sigma$) increase of the inferred astrophysical normalization. We thus consider the astrophysical fit results, presented here, to be robust against significant changes in the hole-ice modelling. Besides the crucial impact of the hole-ice model on the fit quality, one might wonder how the quality of the fit is further influenced by the remaining four systematic parameters and the corresponding constraints from auxiliary data. Assuming the 30 cm hole ice model to represent the true detector response together with all other detector-related systematic parameters to kept fixed at their nominal values, we would still find an acceptable goodness of fit p-value of 39% (this scenario corresponds to the green contour in Fig. 6.9, bottom left). This question can be re-framed by asking whether the dataset in this work (together with the auxiliary datasets) provides evidence for further departures from the nominal detector and flux models beyond the change in hole-ice scattering. In other words, is this neutrino dataset inconsistent with the calibration measurements? The corresponding likelihood ratio test, which requires comparison to a fit without auxiliary constraints, yields $p = 0.55$ and thus is not significant[3]. This result of course would not justify to remove the "insignificant" parameters from the model. As argued in Sec. 5.7 they are needed to realistically model the corresponding systematic uncertainties.

We finally studied whether a different statistical treatment of the nuisance parameters would have led us to different conclusions. This is done by repeating the analysis using Bayesian methods. In particular we now interpret the constraints derived from auxiliary measurements as prior information to be encoded via corresponding prior probability distributions. We then calculate the joint posterior distribution of all 10 parameters and arrive at he two-dimensional joint posterior distribution of the astrophysical parameters (Φ, γ) by marginalizing over the remaining 8 nuisance parameters. The 68% (95%) highest posterior density credible regions compare very well to the 68% (95%) confidence

[3]This takes into account the uncertainty in the calibration measurements. We could ignore it and pretend to have perfect calibration measurements. In this case one would test against the nominal values only. The result remains insignificant with $p = 0.07$.

contours from before. The joint posterior distribution for the astrophysical parameters is shown in Fig. 6.9 (bottom right) together with the aforementioned contours.

Uncertainties Related to Hadronic Interactions

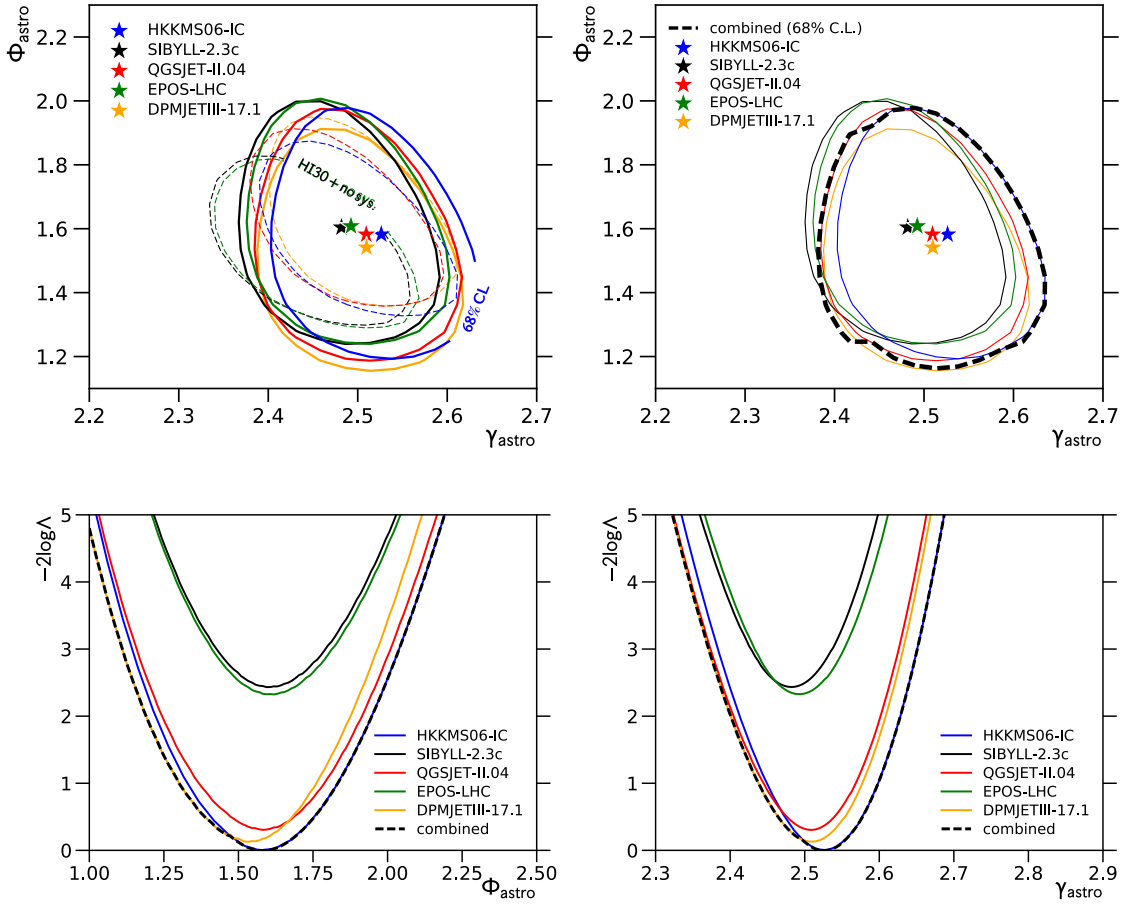


FIGURE 6.10: Left: Impact of different hadronic interaction models on confidence contour - before consideration of detector systematics (dashed) and after (solid). Top right: statistical combination of hadronic interaction uncertainties. Bottom: Table of best-fit values for each alternative interaction model.

We discussed in Sec. 4.2.1 that uncertainties in the modeling of hadronic interactions effect the shape and normalization of the atmospheric conventional neutrino flux prediction, the largest background for our analysis of astrophysical neutrinos. We investigated the sensitivity of our astrophysical flux measurement to the details of the hadronic interaction models by replacing the nominal conventional flux prediction by alternatives calculated with MCEq [169] for the following hadronic interaction models: SIBYLL 2.3c, EPOS-LHC, QGSJETII v04 and DPMJETIII 17.1 (c.f. Sec. 5.7). The results are given in Tab. 6.2. The best-fit values for the astrophysical spectral index range from $\gamma = 2.48$ (SIBYLL 2.3c) to $\gamma = 2.51$ (DPMJETIII-17.1). All alternative hadronic models considered yield insignificantly harder spectral indices than the nominal one. The quality of

Parameter		Hadronic Model	Result $\pm 1\sigma$	$-2\log\Lambda_{sat}$
spectral index	γ		$\mathbf{2.51^{+0.07}_{-0.08}}$	
norm astro	ϕ		$(\mathbf{1.54^{+0.24}_{-0.26}})$ c.u.	
cosmic ray index dev.	$\Delta\gamma_{CR}$		0.01 ± 0.03	
norm conv	ϕ_{conv}	DPMJETIII-17.1	$(1.15^{+0.14}_{-0.13}) \times \Phi_{MCEq}$	65.2
spectral index	γ		$\mathbf{2.49 \pm 0.07}$	
norm astro	ϕ		$(\mathbf{1.61^{+0.26}_{-0.22}})$ c.u.	
cosmic ray index dev.	$\Delta\gamma_{CR}$		0.05 ± 0.03	
norm conv	ϕ_{conv}	EPOS-LHC	$(0.93^{+0.12}_{-0.10}) \times \Phi_{MCEq}$	67.4
spectral index	γ		$\mathbf{2.51^{+0.07}_{-0.08}}$	
norm astro	ϕ		$(\mathbf{1.58 \pm 0.25})$ c.u.	
cosmic ray index dev.	$\Delta\gamma_{CR}$		0.04 ± 0.03	
norm conv	ϕ_{conv}	QGSJETII-04	$(0.99^{+0.12}_{-0.11}) \times \Phi_{MCEq}$	65.3
spectral index	γ		$\mathbf{2.48 \pm 0.07}$	
norm astro	ϕ		$(\mathbf{1.60^{+0.26}_{-0.22}})$ c.u.	
cosmic ray index dev.	$\Delta\gamma_{CR}$		0.05 ± 0.03	
norm conv	ϕ_{conv}	SIBYLL-2.3c	$(1.02^{+0.13}_{-0.11}) \times \Phi_{MCEq}$	67.5

TABLE 6.2: Table of best-fit values for each alternative interaction model.

the fit for QGSJETII v0.4 and DPMJETIII-17.1 is very similar to the default result presented earlier in this Chapter, while the predictions from SIBYLL-2.3c and EPOS-LHC lead to larger values for $-2\log\Lambda_{sat}$ by about 2.5 units. This difference is small compared to the standard deviation ($\sim$ 12 units) of the chi-squared distribution that governs the goodness-of fit measure and thus all models lead to acceptable fits. Fig. 6.2 (top left) shows the two-dimensional 68% confidence contours for the astrophysical flux parameters calculated for each model (solid lines). Also shown are the contours one would obtain if systematic detector uncertainties are not included in the fit (dashed). In the latter case, different hadronic interaction models would allow for a very noticeable shift of the contour towards harder spectra with lower normalization. It appears less pronounced when detector systematics, which effectively reduce the sensitivity of the fit to the differences between the models, are accounted for. The combined contour is shown in Fig. 6.2 (top right) as dashed black line. It has been derived via discrete profiling over the different results (colors), as discussed in Sec. 5.7.2. Corresponding combined 1d profile likelihood functions are shown at the bottom (left: normalization, right: spectral index). These determine the final uncertainties given in eqs. (6.1) and (6.2) and are slightly larger than the ones given in Tab. 6.1 assuming the HKKMS06 flux prediction [99].

6.2 Comparisons with other IceCube Measurements

Searches for high-energy astrophysical neutrinos have a long history in IceCube (and its predecessor AMANDA). Traditionally two search strategies are employed. Event selections focusing on muon neutrinos from the Northern Sky are essentially free of atmospheric muon background since the Earth serves as an impenetrable shield against muons

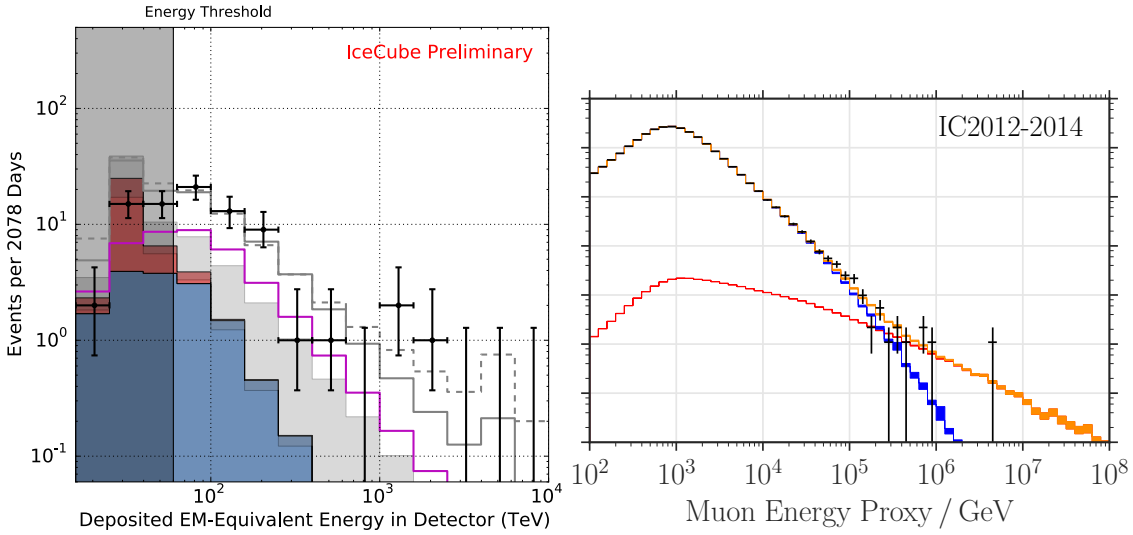


FIGURE 6.11: Measurements of the high energy, astrophysical neutrino flux with the IceCube detector. High Energy Starting Events (HESE 6-year [204], left) and Northern Sky Muon Neutrino Track Events (Tracks 8-year [205]).

[206][207][201]. In addition they achieve large effective areas at highest energies, since muons from neutrino interactions far outside the detection volume can be observed. However, due to the large background from atmospheric muon neutrinos, the observation of astrophysical neutrinos is limited to the $E \gtrsim 100\,\text{TeV}$ energy range. Event selections based on cascades instead are sensitive to lower energies and events from all directions but are statistically limited at highest energies $E \gtrsim 100\,\text{TeV}$ [208][209][155]. A related idea is to search for starting events, i.e. cascades and starting tracks that are contained within the detection volume. The outer part of the detector is used as a veto against incoming muon background. The latter strategy provided the first evidence for the existence of high energy astrophysical neutrinos [5]. The corresponding selection of events has been named HESE (High Energy Starting Events). It is mainly sensitive in a rather narrow energy range ($\sim 100\,\text{TeV}$), since atmospheric muon background becomes important below $E_{rec} \sim 60\,\text{TeV}$ and statistics is limited at much higher energies. The latest observed HESE energy spectrum is shown in Fig. 6.11 (left) and contains 82 events observed in 6 years of data taking. Shortly after the initial discovery of the high-energy neutrino flux in the HESE-channel, a neutrino flux with similar properties was discovered using muon neutrinos from the Northern Sky [207][201], called "tracks" hereafter. The observed spectrum with tracks shown in Fig. 6.11 (right) for the latest years of data taking (2012-2016). Since the original HESE analysis is limited by its relatively high energy threshold, attempts have been made to lower it. Two such analyses have been developed in parallel using two years of data (2010/11): Medium Energy Starting Events (MESE) [164] (Fig. 6.12 top) and a Cascades [155] (Fig. 6.12 2nd row). The measured spectral indices fall between the ones measured with HESE and tracks. In addition a combined analysis of both channels (tracks and starting events) has been performed using the first two years of data (2010/11). The astrophysical flux was found to be consistent with an equal mixture of all neutrino flavors, as shown in Fig. 6.12 (bottom right). Searches for point sources of neutrino emission in the sky or correlations of these events with the

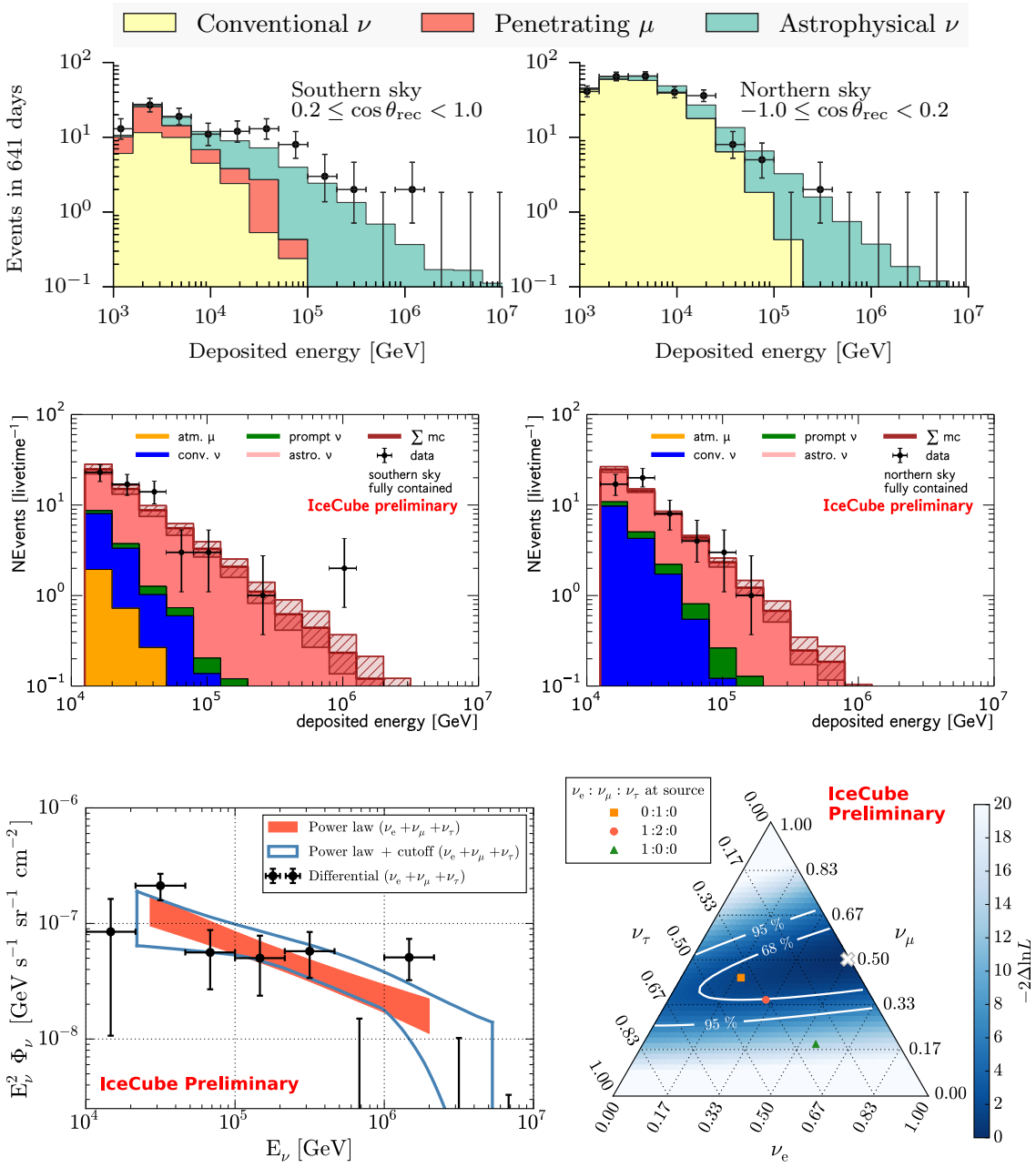


FIGURE 6.12: Measurements of the high-energy, astrophysical neutrino flux with the IceCube detector. Medium Energy Starting Events (MESE 2-year [164], top row), Cascade Events (Cascades 2-year [155], 2nd row) and a combined fit to several samples (Global Fit 2-years [199], bottom row). Measurement of flavor composition by [199] (bottom right).

galactic center/plane have not obtained significant results and limits were set [62][87].

Fig. 6.13 shows the most recent results concerning the two astrophysical parameters, Φ and γ, obtained by various IceCube analyses in comparison to the result presented in this dissertation (blue contour) assuming the astrophysical neutrino flux to behave like a single powerlaw. The different results are from the previous cascade-based flux measurement [155] (Cascades-10/11, pink contour), the most recent High Energy Starting Event analysis [204] (HESE-6y, green contour), the first Medium Energy Starting Event analysis [164] (MESE-10/11, black contour) and the most recent measurement using muon neutrino induced track events from the northern sky [205] (tracks-8y, red contour). Our result, leading in precision, is statistically independent from all previous measurements, except HESE-6y, either because of a different data taking period (Cascades-10/11, MESE10/11) or because a different event signature (tracks-8y) has been analyzed. While cascade analyses are most sensitive in the few tens of TeV energy region, the track based analysis observes astrophysical neutrinos at higher energies above 119 TeV [205]. The different results shown in Fig. 6.13 range from hard spectra ($\gamma = 2.19 \pm 0.10$) with low normalization ($\Phi = 1.01^{+0.26}_{-0.23}$ c.u.), preferred by track-like events, to very soft spectra ($\gamma = 2.92^{+0.29}_{-0.33}$) with high normalization ($\Phi = 2.46 \pm 0.8$ c.u.) obtained by HESE6y, where we defined $1\,\text{c.u.} \equiv 10^{-18}\,\text{GeV}^{-1}\text{s}^{-1}\text{sr}^{-1}\text{cm}^{-2}$. The result obtained in this work falls right in between. The general trend is, as first pointed out in [155] and [201], that cascade and starting event based measurements with lower energy thresholds than the track based measurements appear to prefer softer spectra with higher normalization. Hence the question of consistency arises. We developed a rather general and intuitive method of comparing independent multivariate measurement results [155] and testing them for consistency[4], which we later refined in collaboration with Sebastian Schoenen (one of the corresponding authors of [201]) [210]. We would like to make the following two observations:

1. The result of this work is **well consistent** with previous measurements in a similar energy range and signature (MESE-10/11 and Cascades-10/11) but significantly improves upon the achieved precision.

2. The result of this work appears in **$2.0\sigma - 2.3\sigma$ tension** with the result obtained using muon neutrinos from the northern sky, sensitive at higher energies.

The latter deserves further attention. Fig. 6.14 (top) compares the reconstructed energy spectrum predicted for our sample, assuming the best-fit obtained with muon neutrinos is true (purple), to the best-fit obtained in this work (red). The spectra are very similar, essentially indistinguishable, in the ~ 200 TeV to 1 PeV energy range. They are, however, very different at energies below 200 TeV, where the muon-neutrino based spectrum significantly under-predicts the observed cascade data. At the high energy end, this sample (unfortunately) is limited by the available data statistics and above reconstructed energies of 160 TeV ($= 10^{5.2}$ GeV) only 7 events are observed. By themselves, these 7 events can not provide meaningful constraints on the shape of the spectrum. In particular they cannot distinguish between hard spectra with small normalization and soft spectra with large normalization, exactly the direction along which IceCube results appear to differ

[4] Technically, we only test for evidence of inconsistency - or absence thereof.

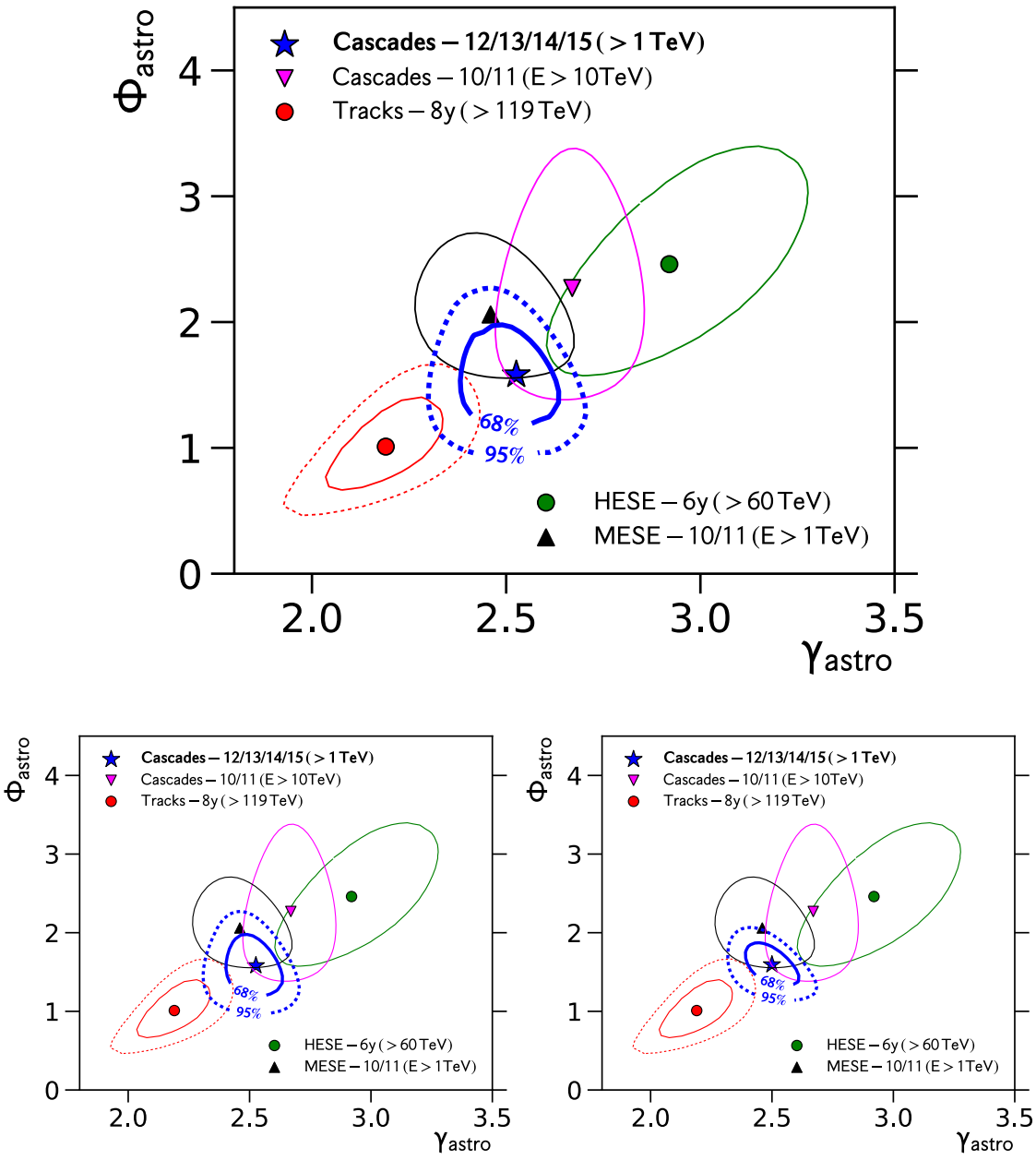


FIGURE 6.13: Top: Comparison of the result obtained here to previous IceCube measurements: ν_μ-tracks (8y) [205] (red), Cascades (2y) [155] (pink), HESE (6y) [204] (green) and MESE (2y) [164]. Bottom left: same as top, but excluding hadronic interaction uncertainties. Bottom right: same as bottom left, but excluding detector systematic uncertainties.

(c.f. Fig. 6.13). This is explored in Fig. 6.14 (bottom left), showing the likelihood as function of the two relevant parameters (Φ, γ) when only those 7 events are considered. The high energy end, especially the Glashow Resonance at 6.3 PeV and the non-observation of corresponding events, requires the normalization to decrease when harder spectral shapes are tested. Since the low energy region is not included in this calculation, soft spectral

indices with large normalization are also in agreement with these events. Fig. 6.14 (bottom left) also shows the contours from the track based (dark blue) and previous cascade based (red) measurements. Independent of the spectral index, the 7 events prefer normalizations smaller than the previous measurements. This under-fluctuation, however, is not significant. Taking into account the uncertainties in the flux measurement using muon

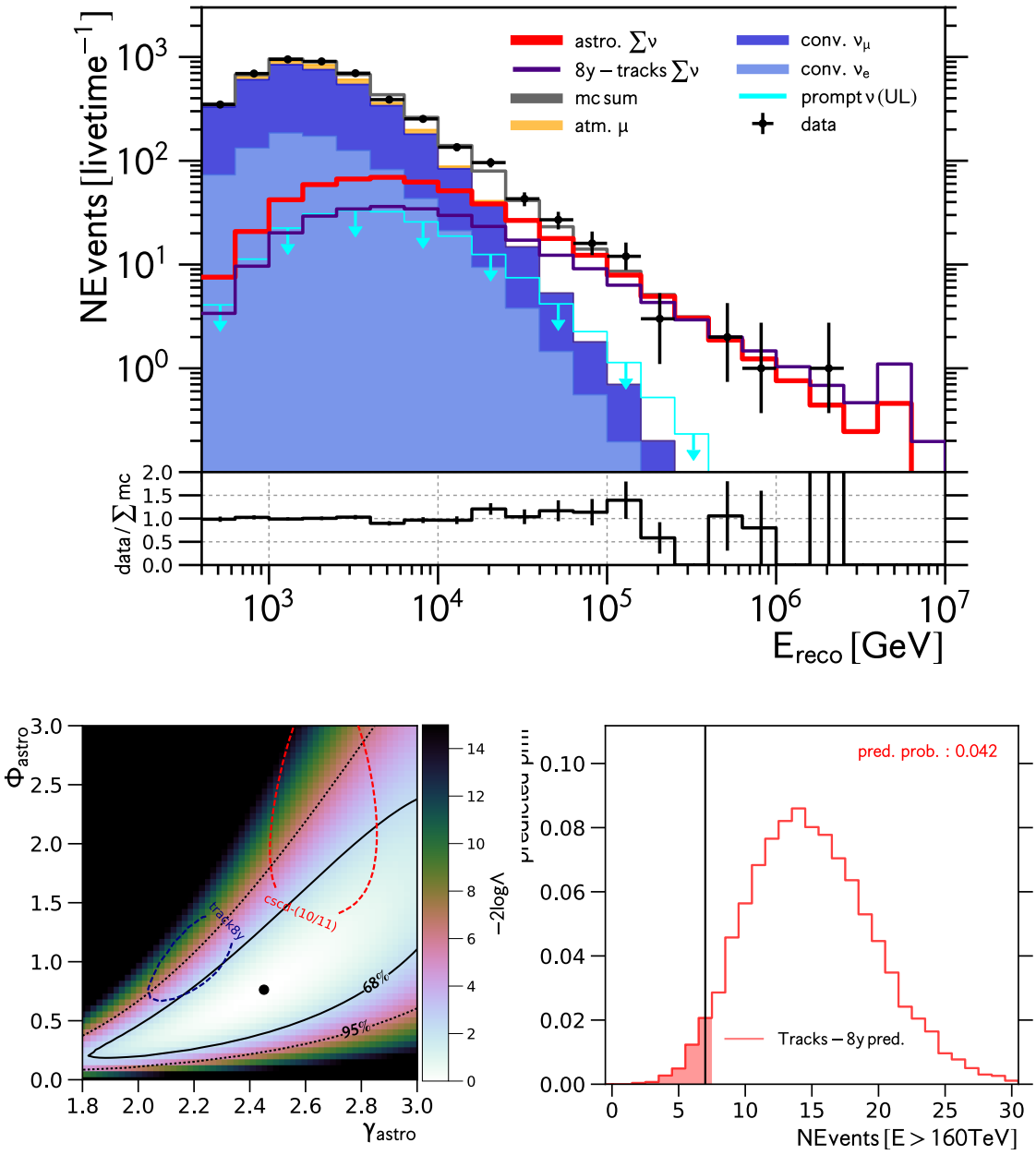


FIGURE 6.14: Top: comparison of the astrophysical contribution assuming the flux measured here (red) with contribution expected from best-fit using ν_μ-tracks [205] (purple). Bottom: confidence contour obtained in this work if only $E_{rec} > 160\,\text{TeV}$ events are considered (left). Posterior predictive distribution of number of expected cascades with $E_{rec} > 160, \text{TeV}$ (right).

neutrinos, we estimate a corresponding posterior predictive probability for observing 7 events or less in this cascade sample with energies larger than 160 TeV of 4%, see Fig. 6.14 (bottom right). We will demonstrate in Sec. 6.6.4 that modeling the astrophysical neutrino flux with a 2-component powerlaw could potentially alleviate some of the tension between both measurements. However to what exact extent this would improve the consistency depends crucially on the ability of the lower energy muon neutrino data to permit a powerlaw component with soft spectral index and small normalization. Hence no firm conclusion can be drawn from our analysis (and other IceCube analyses) alone. One could also be tempted to extend the astrophysical neutrino flux model to be anisotropic in order to explain the discrepancy and, in particular, introduce a north-south asymmetry. The track based measurement is only sensitive to the Northern Sky while the astrophysical neutrinos in this sample, despite coming from all directions, are dominated by the ones from the southern sky. We will show in Sec. 6.6.6 that this only marginally improves the situation due to the increased uncertainties caused by working with a more complex model. The data obtained in our analysis prefers the fluxes in both hemispheres to be identical. Finally it might be worth pointing out that the track measurement at high energies is most sensitive to near-horizontal directions. Hence possible model extensions related to the angular dependence of the astrophysical flux would require additional fine-tuning beyond a simple north/south-asymmetry.

Consistency with HESE-6y

Another question concerns the compatibility between the results obtained in this work and the one obtained in the 6-year High Energy Starting Event measurement [204] (green contour in Fig. 6.13). The preferred spectral index of $\gamma = 2.92$ in that analysis is in significant tension with the low energy data observed here. Superficially one might refer to the large uncertainties in that measurement (large contour) caused by the high-energy threshold of 60 TeV and correspondingly the small number of events observed. A rigorous statistical comparison, between both results, however, is complicated due to the large overlap between both samples. $\sim$ 40% of the events in this sample with reconstructed energies above 60 TeV also contribute to the HESE-6y sample. We will show in Sec. 6.6.3 that for a broken power-law astrophysical neutrino flux, eq. (5.6), we find a best-fit break energy of 70 TeV with spectral index $\gamma_2 = 2.92$ for energies larger than the break energy. This perfectly reconciles both measurements in the energy region that HESE-6y is sensitive to - as expected. The difference lies in the interpretation. Since this sample does not provide significant evidence for the broken powerlaw model over the single powerlaw model, we consider the behavior above 70 TeV as local statistical effect related to the under-fluctuation in the total number of events with energies larger than $\sim$ 160 TeV mentioned above. Of course other interpretations are possible and will be discussed in greater detail in Sec. 6.6.

6.3 Significance: Rejecting Purely Atmospheric Origin

In this section we will evaluate the statistical significance of the observed astrophysical flux in this sample over the atmospheric origin. This is interesting first, because this

Parameter		Calibration	Best-Fit
spectral index	γ	-	–
norm astro	ϕ	-	–
norm conv	ϕ_{conv}	-	$0.93 \times \Phi_{HKKMS06}$
norm prompt	ϕ_{prompt}	-	$\mathbf{15.8 \times \Phi_{BERSS}}$
norm muon	ϕ_{muon}	-	1.37
cosmic ray index dev.	$\Delta\gamma_{CR}$	0.00 ± 0.05	0.06
scattering scale (bulk)	ϵ_{scat}	1.00 ± 0.07(*)	1.08
absorption scale (bulk)	ϵ_{abs}	1.00 ± 0.07(*)	0.98
scattering scale (HI)	ϵ_{abs}	1.00 ± 0.67	1.08
DOM efficiency	ϵ_{eff}	0.99 ± 0.10	1.70
$-2 \log \Lambda_{sat}$	146.36		

TABLE 6.3: Background-only best-fit values.

dataset is statistically independent from the HESE-2y dataset [5], and second, because we significantly improved the treatment of systematic uncertainties for example related to the detector response to light signals. We calculate the significance using the following likelihood ratio hypothesis test (c.f. Sec. 5.4):

$$H_0 : \Phi_{astro} = 0 \quad \text{against} \quad H_1 : \Phi_{astro} > 0 \tag{6.3}$$

This test requires to first fit the background-only hypothesis to the observed data. The corresponding result is given in Tab. 6.3. Without a contribution from astrophysical neutrinos, this sample would require a very large prompt atmospheric neutrino flux of $\Phi_{prompt} = 15.8 \cdot \Phi_{BERSS}$, more than one order of magnitude larger than the nominal flux prediction from [102], and in violation of several upper-limits obtained by IceCube in the past, e.g. [164][201]. Additionally, in order to improve the model fit in the southern sky, increased bulk ($\alpha^{BI}_{scat} = 1.08$) and hole ($\alpha^{HI}_{scat} = 1.70$) ice scattering would be necessary, effectively reducing the impact of the atmospheric neutrino self-veto effect. Still, the background-only model does not describe the data, especially at high energies in the southern sky, as shown in Fig. 6.15. In the northern sky (Fig. 6.15, second row, right) the mismatch is less obvious and appears mostly at intermediate energies ($< 60\,\text{TeV}$) because the prompt atmospheric neutrino spectrum is slightly softer ($\gamma_{prompt} \sim 2.7$) than the preferred value for the astrophysical component of $\gamma_{astro} \sim 2.5$. Overall, we find a observed test-statistic value of $-2 \log \Lambda = 81.3$. As discussed in Sec. 5.4, calculating the corresponding p-value requires knowledge of the relevant distribution of the test-statistic. Wilk's theorem [191] does not apply for two reasons. First, the parameter of interest is tested at the boundary of its parameter space ($\Phi_{astro} = 0$) and second, the spectral index γ constitutes a free parameter that is undefined, if the null-hypothesis was true. Thus, the distribution of the test-statistic is unknown and needs to be estimated from simulating toy-experiments. This was done assuming that the values given in Tab. 6.3 represent the true values realized in nature. Fig. 6.16 (top left) shows the resulting distribution obtained from 4200 toy experiments (red histogram) compared to the observed value (vertical black line). Many orders of magnitude more toy-experiments would be needed to reliably estimate the tail-area probability of observing $-2 \log \Lambda > 81$ (red shaded area). The black dashed line shows one possible, ad-hoc extrapolation of our

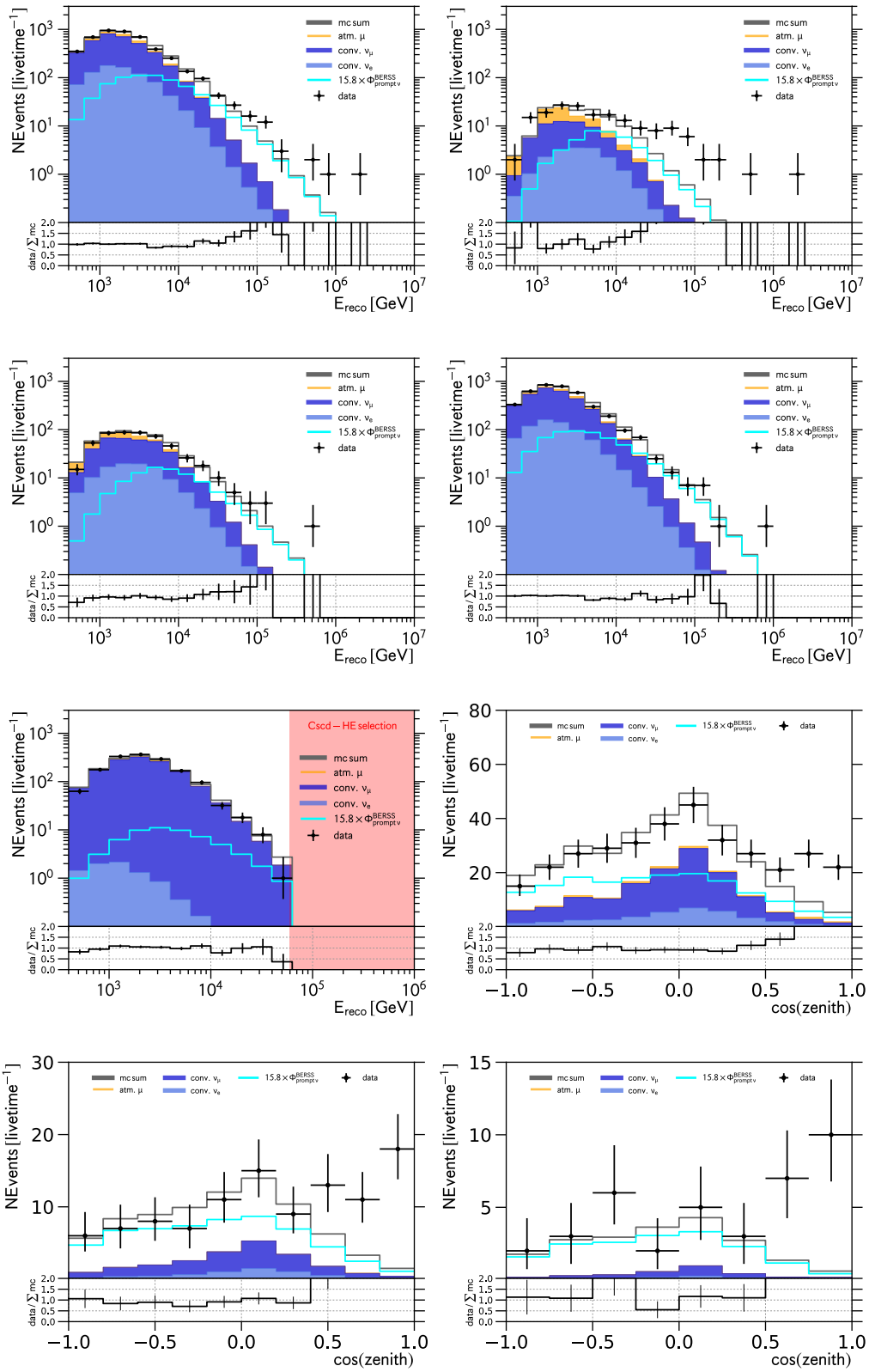


FIGURE 6.15: Best-fit distributions assuming no contribution from astrophysical neutrinos $\Phi_{astro} = 0$. Cascades from all directions (top left), Cascades with $0.6 < \cos\theta_{rec}$ (top right), Cascades with $0.2 < \cos\theta_{rec} \leq 0.6$ (middle left), Cascades with $\cos\theta_{rec} \leq 0.2$ (middle right). Energy spectrum of starting tracks (3rd row, left). Zenith distribution of cascades for $E_{rec} > 10\,\text{TeV}$ (3rd row, right), Zenith distribution of cascades for $E_{rec} > 25\,\text{TeV}$ (bottom left) and for cascades with $E_{rec} > 60\,\text{TeV}$ (bottom left).

simulations. We estimate a tail-area probability of $\sim 10^{-18}$ with large uncertainties due to the mathematically unjustifiable extrapolation and the occurrence of multiple minima in the likelihood function of individual toy-MC datasets which can cause the fit to converge to wrong solutions in individual cases. Clearly, an independent confirmation of this estimate is needed.

Notice that if the spectral index γ was known ahead of the experiment, it would not constitute a free parameter that is undefined under the null-hypothesis. Instead it would just be one constant. In this case Wilk's theorem is still not applicable, because of the parameter boundary ($\Phi_{astro} = 0$). However, a theorem derived by Chernoff [193] is. For our case with $\gamma = c$ (c being some constant) it follows asymptotically

$$\begin{aligned} -2\log\Lambda(\mathbf{x}) &\xrightarrow{d} f_\xi(\xi) \\ f_\xi(\xi) &= \frac{1}{2} \times I(\xi) + \chi^2_1(\xi) \\ I(\xi) &= \begin{cases} 1 & \xi = 0 \\ 0 & \xi > 0 \end{cases} \end{aligned} \tag{6.4}$$

We confirmed that eq. (6.4) applies to our measurements by means of simulation ($\gamma \in \{2.0, 2.5, 3.0\}$). The result is shown in Fig. 6.16 (top right). For spectral indices that resemble the spectral shape of the prompt atmospheric flux ~ 2.7 the prediction from eq. (6.4) matches our simulation well. For $\gamma = 2.0$ the asymptotic expectation over-predicts the tail behavior, which for our purpose appears conservative.

We can now relate our search for a non-zero astrophysical flux with free, variable spectral index to a search with fixed spectral index (see above) thanks to a result by Davies [211]. The idea is to bound the p-value from above by calculating the *local* p-value expected from a search with fixed spectral index and then add a correction for the so-called *look elsewhere effect* [212] that changes the local p-value into a *global* one. The procedure is well described in [212] and further advocated by different authors in [213] and [214]. The global p-value is approximated by

$$p = \frac{P(\chi^2_1 > \xi_{obs})}{2} + E[U(c_0)|H_0] \times \exp\left(-\frac{\xi_{obs} - c_0}{2}\right) \tag{6.5}$$

where $E[U(c_0)|H_0]$ is the expected number of up-crossings $U(c_0)$ of the corresponding $\frac{1}{2}\chi^2_1 + \delta(0)$-random process [213], assuming H_0 is true, and c_0 is an arbitrary threshold with $c_0 << \xi_{obs}$. Since c_0 can be chosen small (we use $c_0 = 0.5$, suggested in [212]), a relatively small set of toy-experiments is sufficient to accurately estimate $E[U(c_0)|H_0]$. We choose a search region $1.0 < \gamma < 4.0$ and implemented the numerical algorithm discussed in [213] using a grid with 100 evaluation points. Despite being vastly more efficient than trying to "brute-force" $O(10^{19})$ toy-fits, the method is also inherently more stable against the occurrence of multiple minima due to the evaluation on a grid that reduces the number of dimensions, the numerical minimization is performed in. The procedure is demonstrated in Fig. 6.16 (2nd row, left) for a few randomly selected toy-datasets, where up-crossings are marked with stars. Fig. 6.16 (2nd row, right) shows the distribution of the number of up-crossings of the test-statistic above $c_0 = 0.5$ that we obtained from a

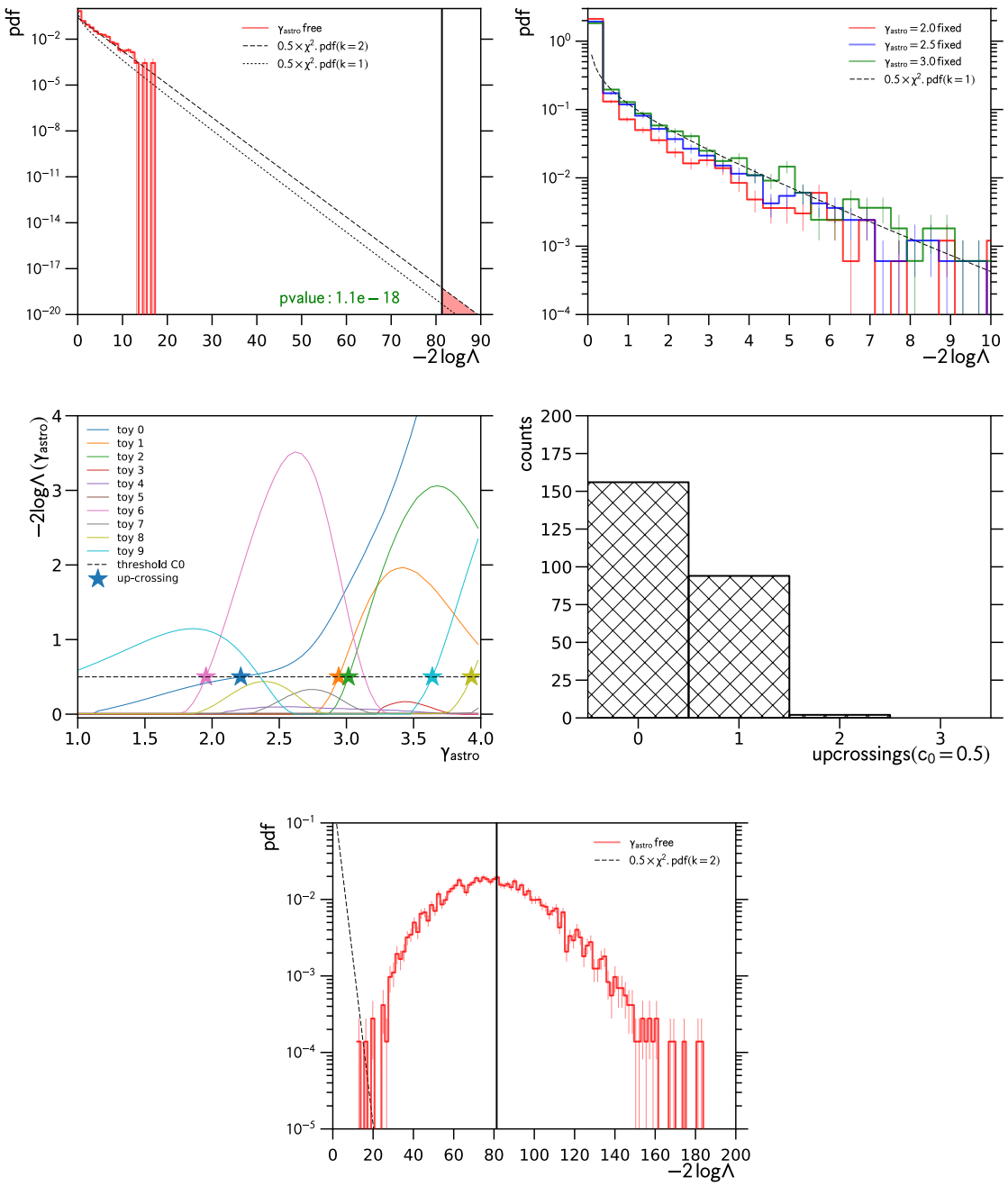


FIGURE 6.16: Calculation of significance over purely atmospheric origin: test-statistic $-2\log\Lambda$ distribution from toy experiments without signal (top left). test-statistic $-2\log\Lambda$ distribution from toy experiments for fixed spectral indices (top right). Examples of up-crossings of LRT-random process (middle left). Distribution of number of up-crossings (middle right). Test-statistic $-2\log\Lambda$ distribution from toy experiments with signal (bottom).

sample of 252 toy experiments. This yields $E[U(c_0 = 0.5)|H_0] = 0.39$. Finally, using eq.

(6.5), we find a global p-value of

$$\mathbf{p = 1.2 \cdot 10^{-18}} \tag{6.6}$$

which compares well to the previous estimate. The result corresponds to a trials factor of 12.4 relative to a search with fixed, pre-specified value for the spectral index γ.

The final result is a preference for a single powerlaw astrophysical neutrino flux (with a-priori unknown intensity and spectral shape) over purely atmospheric neutrino production corresponding to a **significance of 8.7σ**, obtained with a dataset that is statistically independent from the initial evidence reported in [5], and taking into account all relevant systematic uncertainties. Without the inclusion of systematic uncertainties, described in 5.7, the significance would of course be larger at 9.9σ[5]. We effectively more than doubled the significance compared to [5], despite only doubling the total live time of the experiment.

Lastly, we checked whether the high significance appears unusual in a statistical sense. We re-calculated the test-statistic distribution assuming that the true values of our model, including an astrophysical flux component, are given by Tab. 6.1. The resulting distribution of the test-statistic values ($-2\log\Lambda$) is shown in Fig. 6.16 (bottom right). In particular we find that 44% of the replicated toy-experiments would yield a more significant result than our actually observed dataset. From our simulations we derive a 68% central range for the predicted significance of $7.3\sigma - 9.8\sigma$ and thus our result appears well consistent with expectations.

6.4 Energy Range

The "energy range" of this measurement serves an important purpose in the interpretation of atmospheric and astrophysical flux measurements. If no signal is observed and therefore upper-limits are obtained, the energy range conveys additional information about the energies at which the corresponding flux with intensity above the limit would be excluded. Obviously, if for example from theory one expects the flux to be present also below the detection threshold energy of the detector, the upper-limit does not apply to that energy region. Similar considerations apply if a flux is observed. Models often extend beyond the detection sensitivity of the experiment, while observations can for obvious reasons only be performed above threshold. While this reasoning appears immediately intuitive, how to formalize it is less clear. Various definitions and methods have been devised and applied over time within the IceCube Collaboration, for example the different approaches discussed in [200][164][205]. When applied to this work the question to answer is: within what range in neutrino energy does the data support the existence of a single-powerlaw astrophysical neutrino flux? We will address this question by first deriving the energy range using the definition put forward in [205] before introducing yet one more approach. In Sec. 6.3 we discussed the calculation of the significance of the observed astrophysical

[5] We did not repeat all toy-simulations, since we do not anticipate any change in the distributions of the test-statistic. We only calculated the observed TS-value corresponding to the nominal detector except for a (fixed) Hole-Ice scattering length of 30 cm.

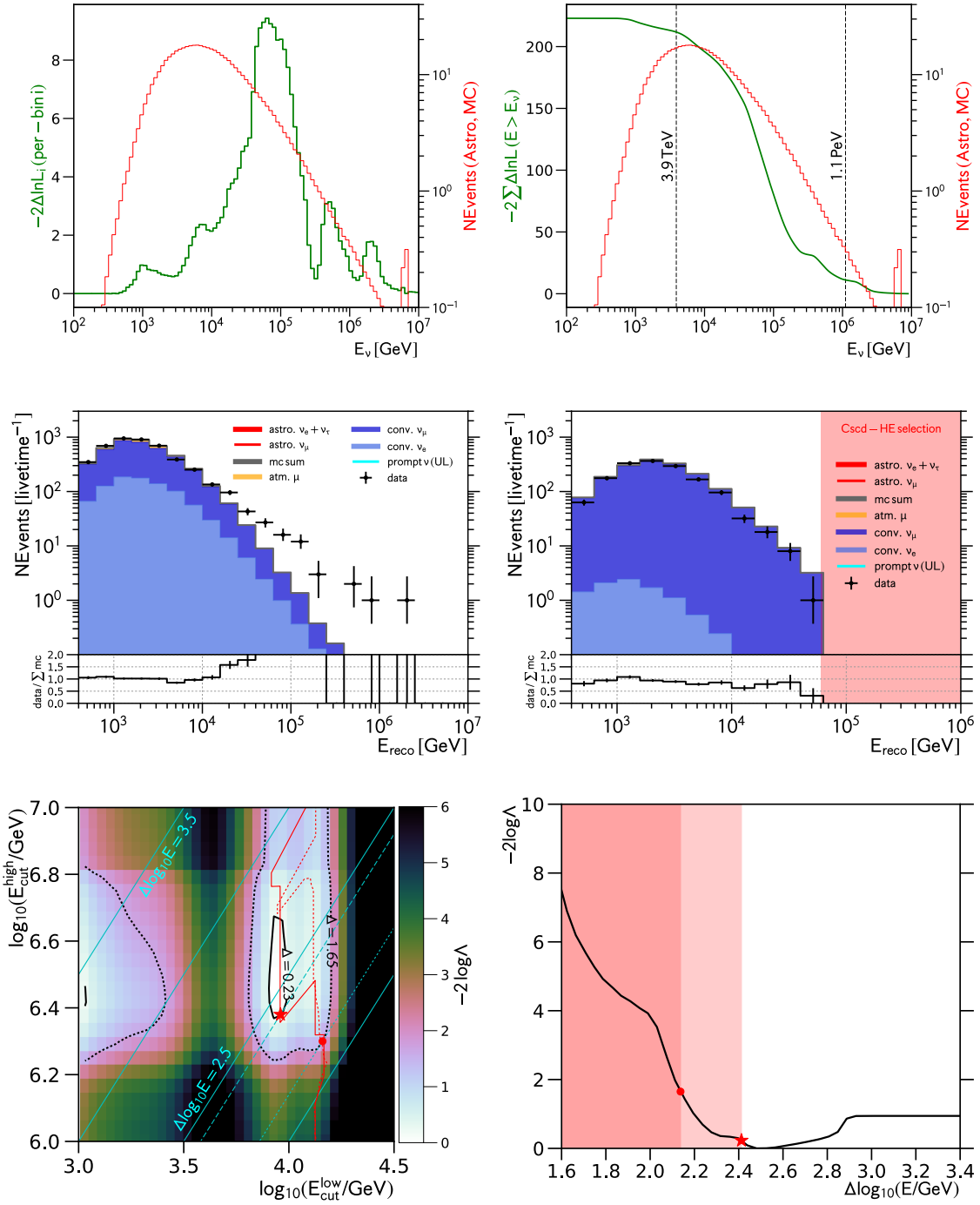


FIGURE 6.17: Calculation of valid energy range: likelihood contribution in each bin (top left). summed likelihood contribution from bins with $E' > E$ (top right). Middle: Cascade (left) and starting track (right) energy spectra assuming zero contribution from prompt atmospheric and astrophysical neutrinos. Bottom: construction of lower-limit on logarithmic range of flux (see text).

flux over the atmospheric background only hypothesis, especially the central role of the test-statistic $\xi = -2 \log \Lambda$. The idea behind the construction of the energy range is to estimate a range of neutrino energies that contributes a large fraction (90%) to the absolute value of ξ. A slightly modified definition for ξ is used, that corresponds to a hypothesis test with known atmospheric prompt flux (equal to its best-fit value). The log-likelihood function, eq. (5.18), and thus the test statistic ξ, is defined as a sum over all observable bins. One can therefore easily identify the contribution to ξ from each individual observable bin i, denoted ξ_i. Each observable bin also has a non-zero contribution from the astrophysical neutrino flux. Assuming the best-fit astrophysical neutrino flux and using the standard MC-simulations it is straightforward to calculate the normalized distribution of neutrino energies, $f\left(E_\nu \,|\, i,\, \Phi,\, \gamma\right)$, that contribute to said observable bin. One can then assign a weight ξ_i to this distribution. The weighed sum of all this distributions over all observable bins then could be interpreted as the distribution of the test-statistic ξ in neutrino energy space. If we assume K neutrino energy bins, this reads:

$$\tilde{\xi}_k \left(E_\nu^k\right) = \int_{\Delta E_\nu^k} \sum_i f\left(E_\nu \,|\, i,\, \Phi,\, \gamma\right) \times \xi_i \tag{6.7}$$

A detailed explanation of this method can be found in [215] (Appendix G). The result of this construction when applied to this dataset is shown in Fig. 6.17 (top left, green). One can clearly see that he bulk of the total test-statistic ξ is estimated to stem from between few tens of TeV to few hundreds of TeV. Isolated peaks are visible at highest energies, corresponding to individual events that contribute strongly to the test-statistic. This strong contribution from individual events shows that highly energetic neutrinos provide a lot of evidence for an astrophysical origin. However, the are not very informative about the spectral properties of this flux, due to their small numbers (c.f. Sec. 6.2). The sum of the contributions from neutrino energy bins above some threshold energy E_ν is shown in Fig. 6.17 (top right, green). Using this construction, we estimate that 90% of the total test-statistic $\xi = -2 \log \Lambda$ stem from neutrino energies ranging from

$$E_{low} = 3.9\,\text{TeV} \quad \text{to} \quad E_{high} = 1.1\,\text{PeV}. \tag{6.8}$$

Fig. 6.17 (2nd row) shows the observable distributions for the best-fit in a scenario that has no contribution from either prompt atmospheric or astrophysical neutrinos, for cascade-like events from all directions (left) and starting tracks (right). They should be compared to their best-fit counterparts that include an astrophysical component (Fig. 6.1 and Fig. 6.3). One notices that in the former scenario, the fit adjusts the systematics parameters in order to match the statistically dominant cascade sample at low energies ($E < 10\,\text{TeV}$). This adjustment, while improving the overall fit-quality, comes at the expense of a reduced fit in the starting track sample. Thus, for the bins below 10 TeV the main contribution to ξ is actually obtained from the starting track control sample. However, as shown in Fig. 6.3 (bottom left), the contribution of astrophysical neutrinos, estimated in the latter scenario, is negligible compared to the conventional background. For this reason it does not appear sensible to "assign" the corresponding contribution to ξ to astrophysical neutrinos and to then estimate the neutrino energy range using the prescription described above, including eq. (6.7) - at least for the dataset obtained in this work.

Because of this conceptual difficulty, we devised an alternative calculation. We attempt to answer the following question: what are the minimum and maximum energies (E_ν^{min} and E_ν^{max}) corresponding to the 68% (90%) C.L. lower bound on the logarithmic width of the corresponding range of neutrino energies $\Delta \log_{10}(E_\nu/\mathrm{GeV})$, assuming that the astrophysical neutrino flux behaves as a single power law within a finite range? In other words, what is the minimal logarithmic range, one could limit the flux with, that remains consistent with the data (at a certain confidence level)? Thus we need to extent the unbroken, single power law baseline model with two parameters to describe the possibility of a finite range, i.e. $\log_{10}(E_\nu^{min}/\mathrm{GeV})$ and $\log_{10}(E_\nu^{max}/\mathrm{GeV})$ corresponding to the "single power law box" model from eq. (5.9). We define:

$$\Delta \log_{10}(E_\nu/\mathrm{GeV}) = \log_{10}(E_\nu^{max}/\mathrm{GeV}) - \log_{10}(E_\nu^{min}/\mathrm{GeV}) \tag{6.9}$$

We discussed methods of interval construction in Sec. 5.4. Here we would like to find a 68% (90%) C.L. lower limit on $\Delta \log_{10}(E_\nu/\mathrm{GeV})$, i.e. construct a one-sided interval. This can be done by inverting the following hypothesis test:

$$H_0 : \Delta \log_{10}(E_\nu/\mathrm{GeV}) = c_0 \quad \text{against} \quad H_1 : \Delta \log_{10}(E_\nu/\mathrm{GeV}) > c_0 \tag{6.10}$$

Fig. 6.17 (bottom left) shows the two-dimensional profile-likelihood function for the two extra parameters, where we treated the remaining astrophysical parameters, Φ and γ, as additional nuisance parameters. Calculating the observed test-statistic ($-2\log\Lambda$) value corresponding to the test in eq. (6.10) requires to minimize that function along the axis, where $\Delta \log_{10}(E_\nu/\mathrm{GeV})$ is constant (light blue lines). The corresponding path is shown as the solid[6] red line. The resulting one-dimensional profile likelihood function for $\Delta \log_{10}(E_\nu/\mathrm{GeV})$ is shown in Fig. 6.17 (bottom right). Since this is a one-sided problem, we invoke Chernoff [193], i.e. use eq. (5.27) and find critical values of $-2\log\Lambda = 0.23$ ($-2\log\Lambda = 1.65$) for 68% (90%) C.L. lower limits. The excluded regions of $\Delta \log_{10}(E_\nu/\mathrm{GeV})$ are marked as red shaded regions, where the corresponding limits are shown as red star ($\Delta \log_{10}(E_\nu/\mathrm{GeV}) = 2.1$ at 68% C.L.) and red dot ($\Delta \log_{10}(E_\nu/\mathrm{GeV}) = 2.4$ at 90% C.L.). The 68% C.L. lower limit correspond to a range of

$$E_{low} = 9.1\,\mathrm{TeV} \quad \text{to} \quad E_{high} = 2.4\,\mathrm{PeV} \tag{6.11}$$

while for the 90% C.L. lower limit we find

$$E_{low} = 14\,\mathrm{TeV} \quad \text{to} \quad E_{high} = 2.0\,\mathrm{PeV} \tag{6.12}$$

The upper ends derived in this way are higher than the one found using the previous approach, eq. (6.8). This is because we can be reasonably certain that there exists an astrophysical flux around 2 PeV, given that we have found one event with a reconstructed energy at that value. At the same time we found the lower ends to move to higher energies as well. This is probably because, using the new method, we treated the astrophysical parameters of interest (Φ, γ) as unknown. Finally we should mention that in our construction we did not treat the detector-related systematics parameters as unknown but

[6] This is actually a non-trivial problem. Straight application of our standard, gradient aware minimizer would have yielded the red dotted (wrong) solution. The minimizer follows the gradient and misses the global minimum.

kept them fixed at their nominal values (except for the scattering length in IceCube's drill holes, where we use the best-fit value of 30 cm) to ease the computations. Note that our approach is similar in spirit to that of [164] but the implementation, statistical justification and thus the interpretation are different.

6.5 Prompt Atmospheric Neutrinos

Assuming the astrophysical neutrino flux to follow an isotropic, single powerlaw this analysis finds no evidence for a prompt atmospheric neutrino flux component in the observed spectrum. The best-fit prompt normalization is vanishingly small, $\Phi_{prompt} = 0 \cdot \Phi_{BERSS}$. We are now interested in deriving an upper-limit on the strength of the flux above which this dataset would disfavor its existence. Fig. 6.18 shows the profile likelihood function of the normalization parameter Φ_{prompt} for the prompt atmospheric neutrinos flux (thick black line). It follows from Wilk's theorem [191] that the 90% C.L. upper-limit Φ^{UL}_{prompt} is given by the (largest) parameter value Φ^{0}_{prompt} at which the observed log-likelihood ratio for testing $H_0 : \Phi_{prompt} = \Phi^{0}_{prompt}$ against $H_1 : \Phi_{prompt} \neq \Phi^{0}_{prompt}$

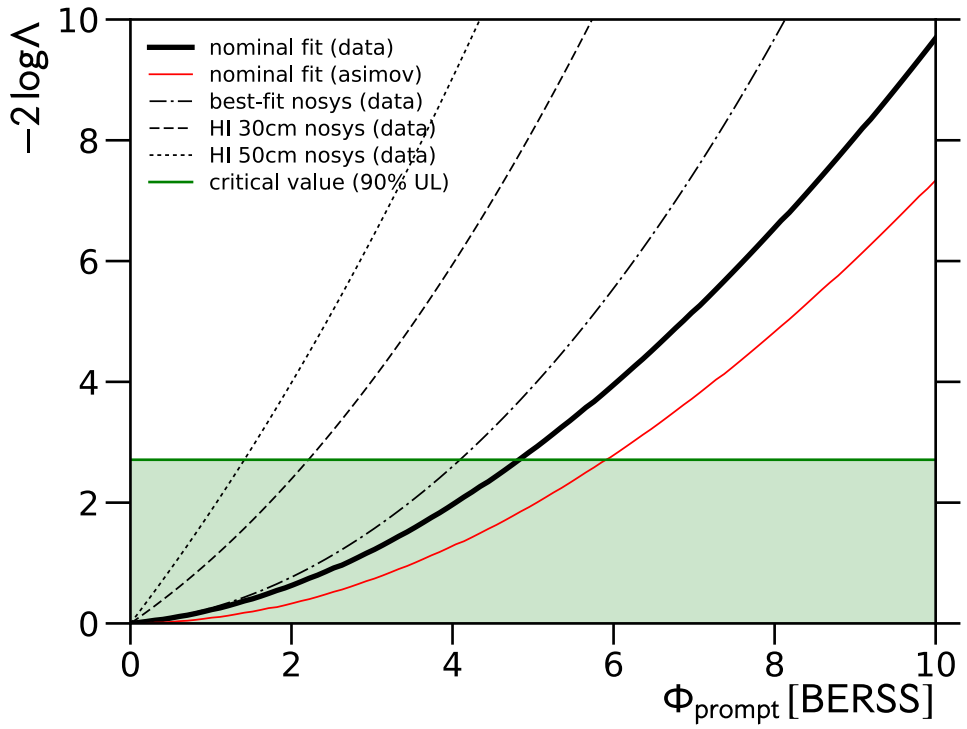


FIGURE 6.18: Calculation of 90% upper-limit on prompt atmospheric neutrino flux. Observed likelihood function including all systematic uncertainties (black) compared to Asimov expectation (red). Different assumptions about detector response (dashed).

attains the value $2\log\Lambda = 2.71$[7]. We find a 90% C.L. upper-limit that is 4.8 times larger than the benchmark prediction from the BERSS model [102].

$$\Phi^{UL90}_{prompt} = 4.8 \cdot \Phi_{BERSS}. \tag{6.13}$$

Tab. 6.4 details the values of all other fit-parameters if the prompt normalization is kept fixed at its upper limit. As expected, in this scenario we would find a slightly harder spectral index of $\gamma = 2.47$ with lower normalization of $\Phi = 1.36$ c.u. compared to the nominal fit result. We also notice a slight increase in hole-ice scattering in this scenario. To confirm that this solution $\Phi^{UL90}_{prompt} = 4.8 \cdot \Phi_{BERSS}$ indeed corresponds to 90% C.L. as inferred from application of Wilks' theorem [191], we calculated the distribution of the relevant test-statistic from toy-experiments assuming the values from Tab. 6.4 as truth. The resulting distribution is shown in Fig. 6.19 and confirms our asymptotic calculation. In particular we determine a confidence level of $(90.6 \pm 0.5)\%$, where the uncertainty is due to the finite size of our simulation ($N = 4200$).
Is our result consistent with expectations? Fig. 6.18 shows the median expected prompt upper limit that one would derive from the profile likelihood function calculated based on the Asimov dataset [192]. This assumes the true value of the prompt normalization to be zero and is called the (median) sensitivity of the experiment. We obtain $\Phi^{UL90exp}_{prompt} = 5.9 \cdot \Phi_{BERSS}$ which is 21% larger than the upper-limit obtained for the real dataset (eq. (6.13)). Setting limits far below the experimental sensitivity can be problematic and be indicative of un-modeled systematic uncertainties or rare fluctuations in the data. Whatever the reason in those cases, the credibility of the result might be questioned [216]. To better understand our result, we studied the distribution of 90% upper-limits, one might have observed, had the data been different. We generate artificial toy datasets and for each calculate the 90% upper limit. In particular we assume

[7] Technically, we are not calculating an upper-limit, but rather a two-sided confidence interval that is bounded from below ($\Phi_{prompt} = 0$). If we had pre-registered, that we want to "measure" an upper-limit, the result would be more restrictive - corresponding to $2\log\Lambda = 1.65$.

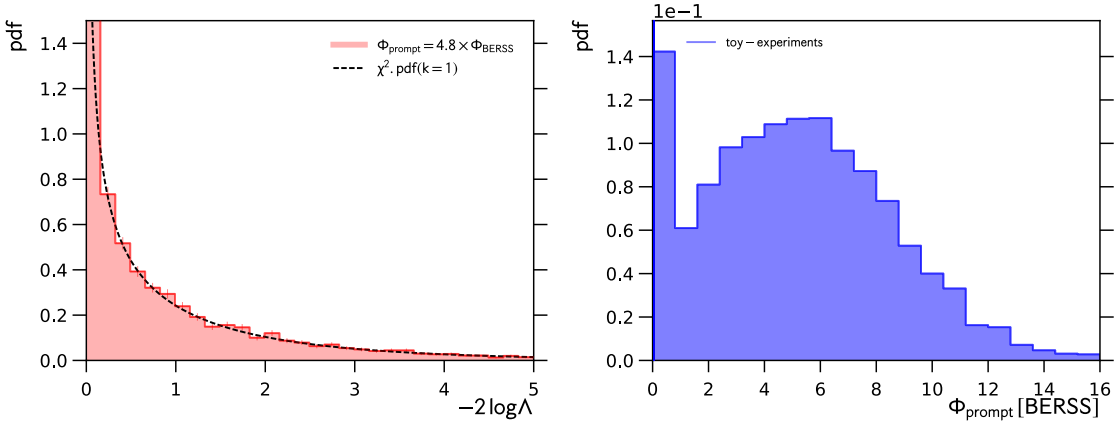


FIGURE 6.19: Left: Distribution of the likelihood-ratio test-statistic assuming a prompt flux at the 90% upper-limit obtained with toy experiments (red) compared to Wilk's expectation (black). Right: Distribution of best-fit values for prompt normalization assuming a prompt flux at the 90% upper-limit.

Parameter		Calibration	Result at $\Phi^{90\%}_{BERSS}$
spectral index	γ	-	2.47
norm astro	ϕ	-	1.36 c.u.
norm conv	ϕ_{conv}	-	$1.05 \times \Phi_{HKKMS06}$
norm prompt	ϕ_{prompt}	-	$\mathbf{4.8 \times \Phi_{BERSS}}$
norm muon	ϕ_{muon}	-	1.41
cosmic ray index dev.	$\Delta\gamma_{CR}$	0.00 ± 0.05	0.05
scattering scale (bulk)	ϵ_{scat}	1.00 ± 0.07(*)	1.03
absorption scale (bulk)	ϵ_{abs}	1.00 ± 0.07(*)	1.00
scattering scale (HI)	ϵ_{abs}	1.00 ± 0.67	1.71
DOM efficiency	ϵ_{eff}	0.99 ± 0.10	1.02

TABLE 6.4: Value for all fit-parameters at the 90% upper limit on the prompt flux. (*) nuisance parameters are anti-correlated with $\rho = -0.1$.

the true value of the prompt atmospheric flux normalization to be zero, corresponding to our best-fit astrophysical flux measurement, Tab. 6.1. The result is shown in Fig. 6.20 (top left) as the blue histogram and covers a rather wide range of possible prompt upper-limits. 68% (95%) of the replicated upper-limits lie within a central interval of $\Phi_{prompt}/\Phi_{BERSS} \in [3.3, 10.2]$ ($\Phi_{prompt}/\Phi_{BERSS} \in [2.1, 14.2]$). These intervals are visualized by the green shaded areas. In particular there is a 36% probability of obtaining a prompt-limit that is more constraining than the one found in this work (eq. (6.13), black vertical line). Thus, our result does not appear unexpected and the small difference, when compared to the median sensitivity, is easily explained by the estimated variance.
From our simulations we can also determine the median 90% upper-limit (green vertical line) which is essentially identical to the asymptotic expectation from the Asimov dataset (red vertical line). Fig. 6.20 also shows the replicated profile-likelihood functions obtained from the toy-experiments (2nd row, left). The Asimov-approximation to the median sensitivity (red) works well, i.e. it overlaps with the values obtained from toy-experiments (green), if values above $\sim 4 \cdot \Phi_{BERSS}$ are tested. Some difference is observed for smaller values and interpreted as a boundary effect that invalidates asymptotic arguments. We have discussed the impact of the same boundary ($\Phi_{prompt} = 0$) on the expected precision of the astrophysical flux measurement in Sec. 6. The same boundary explains the asymmetric shape of the distribution of the replicated 90% upper-limits (Fig. 6.20, top left). Fig. 6.20 (2nd row, right) shows the distribution of the best-fit values for the atmospheric prompt normalization and its pronounced peak at the boundary. Obviously, there exists a correlation between the prompt best-fit value and the corresponding upper-limit. Weaker limits are obtained if the best-fit normalization increases. This is visualized well in Fig. 6.20 (top right). It is plausible to assume that if we were to extent to parameter space to negative values, the distribution of upper-limits would show a heavier tail towards smaller values and thus become more symmetric.

Impact of Systematic Uncertainties
We have found the prompt upper-limit to be quite susceptible to systematic uncertainties. Standard assumptions about the detector response, in particular a hole-ice scattering length of 50 cm would yield a very strong limit of $1.4 \cdot \Phi_{BERSS}$. The corresponding profile-likelihood function is shown in Fig. 6.18 (thin black line, dotted) and increases nearly linearly. Increasing the amount of scattering in the drill holes ($\lambda_{scat} = 30$ cm)

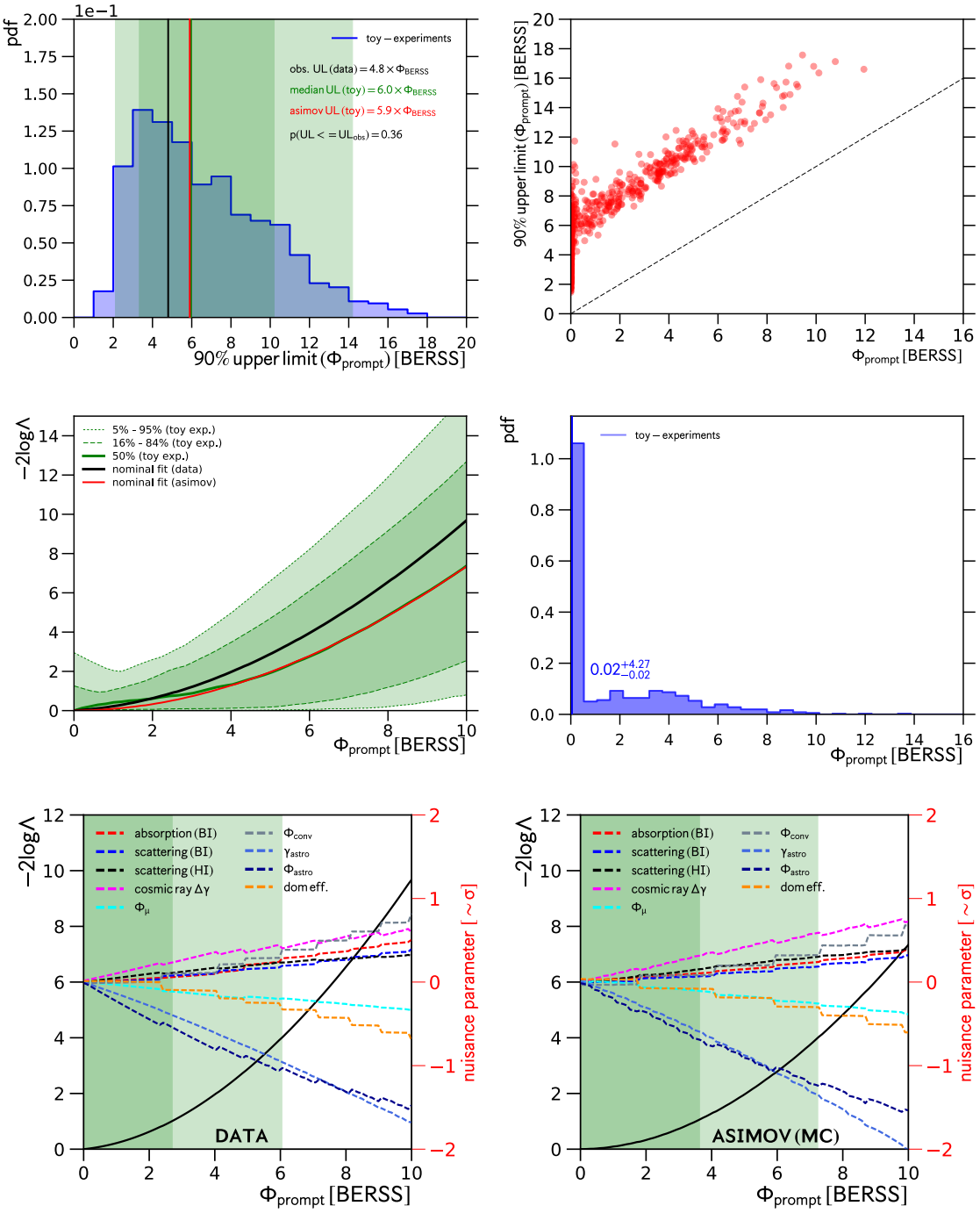


FIGURE 6.20: top left: Sampling distribution of 90% prompt upper-limit assuming no prompt flux exists in nature (sensitivity). Top right: Correlation of 90% prompt upper-limit with best-fit prompt flux normalization. Middle left: Sampling uncertainties on profile likelihood function (green) compared to median expectation from Asimov method (red). Middle Right: Sampling distribution of fitted prompt normalizations. Bottom: Behavior of nuisance parameters as function of inferred prompt normalization.

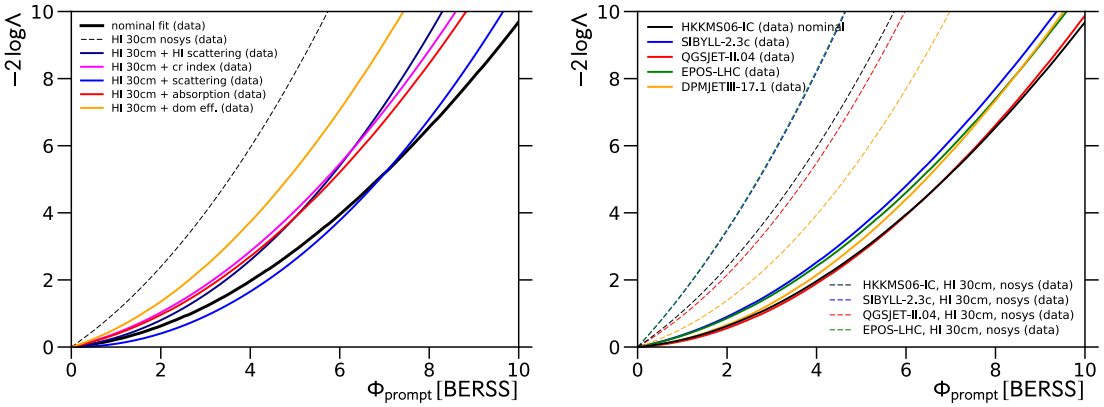


FIGURE 6.21: Impact of systematic uncertainties on prompt flux upper-limit: Left: detector-related systematics. Right: hadronic interaction models

weakens the limit to $2.2 \cdot \Phi_{BERSS}$ (thin black line, dashed), which is still roughly a factor 2 below our final result, eq. (6.13). Adjusting all nuisance parameters to their best-fit values (Tab. 6.1) further weakens the limit to $4.1 \cdot \Phi_{BERSS}$ (thin black line, dash-dotted). The remaining difference is due to the fact that in the final fit, the nuisance parameters are allowed to vary within the constraints provided by this dataset in combination with the auxiliary data. Fig. 6.20 (bottom left) shows the conditional best-fit values for the nuisance parameters as function of increasing normalization of the prompt atmospheric neutrino flux. No single parameter appears to dominate the weakening of the limit and all appear to contribute. This behavior is expected, since it can be reproduced on the Asimov dataset (median expectation), see Fig. 6.20 (bottom right).

It is still insightful to rank the systematics parameter by their individual influence on the prompt limit. Fig. 6.21 (left) shows the profile-likelihood functions for scenarios, in which a (fix) hole-ice scattering length of 30 cm is assumed and one additional nuisance parameter is allowed to vary. This shows that, if judged by their individual impact, the strongest weakening of the limit would be due to uncertainties in bulk-ice scattering. As discussed in Sec. 4.1.2 enhancing bulk-ice scattering biases the reconstructed declination towards the southern hemisphere and thus weakens the observable self-veto effect, and in effect allows for a larger prompt flux.

Finally we studied the effect of hadronic interaction models that can alter the shape of the conventional atmospheric neutrino background. We re-computed the prompt-limit for the different hadronic interaction models, discussed in Sec. 5.7. This is shown in Fig. 6.21 (right), where solid lines are used for the analysis with all systematics being variable, and dashed lines show the results, obtained when all systematics parameters are kept fixed at their nominal values (except for the hole-ice, where we again assume $\lambda_{scat} = 30\,\text{cm}$.). Once all other systematics are accounted for, the different hadronic interaction models only have a minor effect on the limit. In particular, the most conservative limit is given by the default conventional model (HKKMS06) with $\Phi^{UL90}_{prompt} = 4.8 \cdot \Phi_{BERSS}$.

Comparison to previous IceCube Measurements

IceCube has attempted to measure the flux from prompt atmospheric neutrinos in the past. However no analysis has observed a corresponding non-zero contribution to the data

sample. The most recent 90% C. L. upper-limit obtained with a cascade dominated sample has been reported in [164] (Medium Energy Starting Events): $\Phi_{prompt} < 1.52 \cdot \Phi_{ERS}$[8]. A stronger limit has been reported from a measurement using muon neutrinos from the northern sky, $\Phi_{prompt} < 1.06 \cdot \Phi_{ERS}$ [201]. This latter limit is based on a different method than the one presented here. It is conservative in the sense that the method is guaranteed to over-cover (coverage larger than 90%). Using the same method applied in this work, a stronger limit would have been reported: $\Phi_{prompt} < 0.5 \cdot \Phi_{ERS}$, which is a factor ~ 3 smaller than the reported (median) sensitivity of $1.5 \cdot \Phi_{ERS}$ [201]. All limits (including the one reported here) are model dependent in that they assume an isotropic, single power law astrophysical neutrino flux (background for prompt atmospheric neutrinos). Furthermore the experimental (median) sensitivity depends on the assumed true values for the astrophysical neutrino flux. Obviously a better sensitivity is obtained for a hard astrophysical neutrino flux with small normalization than a soft astrophysical neutrino flux with larger normalization. The former predicts less astrophysical background neutrinos at lower energies, where the analyses are most sensitive to a possible signature related to prompt atmospheric neutrinos. According to [215] the (median) sensitivity of the muon neutrino measurement [201] as derived from the Asimov dataset would be weaker than $\sim 4 \cdot \Phi_{ERS}$ if a spectral index of $\gamma_{astro} = 2.5$ with normalization $\Phi_{astro} = 1.7$ was assumed (Fig 12.2 in [215]) - similar to the values preferred by the measurement reported here. This weakening is expected. As the astrophysical spectral index approaches the shape of the prompt flux ~ 2.7 both components become essentially indistinguishable for the muon neutrino based analysis, up to small differences at highest energies to due to CR knee that soften the prompt flux at highest energies. Cascade and Starting Event based measurements are less affected in this scenario due to the atmospheric self-veto suppression of prompt neutrinos in the southern sky. This extra signature, however, is lost if one abandons the assumption that the astrophysical neutrino flux is isotropic and, in particular, introduces a north-south asymmetry (c.f. Sec. 6.6.6). In order to compare these limits to the one reported here, we first have to convert our limit, which is given in units of the BERSS [102] model, into units of the ERS model [101]. In our analysis we find that both models have essentially identical spectral shapes and only differ in normalization, with the letter predicting a larger flux $\Phi_{ERS} \approx 2.82 \cdot \Phi_{BERSS}$. Thus our 90% C.L. upper-limit, eq. (6.13), becomes

$$\Phi_{prompt} < 1.7 \cdot \Phi_{ERS} \tag{6.14}$$

which is larger than both previous results. The corresponding (median) sensitivity becomes $2.1 \cdot \Phi_{ERS}$. Given the discussions above, one can easily understand why this limit is weaker than the one obtained using muon neutrinos from the northern sky [201]. The reason is the preference for a harder astrophysical flux with lower normalization (c.f. Sec. 6.2). If instead both analyses assume an astrophysical flux similar to the one preferred in this work, the sensitivity obtained here would be better by a factor of $2 - 3$, hinging on the assumption that the astrophysical neutrino flux is symmetric with respect to both hemispheres (c.f. Sec. 6.6.6). Assuming an astrophysical neutrino flux with properties given by [201] ($\gamma = 2.13$, $\Phi = 0.9\, c.u.$) the (median) sensitivity of this work would change to $4.1 \cdot \Phi_{BERSS}$ ($1.5 \cdot \Phi_{ERS}$), as determined from the corresponding Asimov

[8] The BERSS [102] prompt flux prediction is an updated calculation compared to the ERS [101] prompt prediction from the same authors. Both calculations derive very similar spectra and differ mostly in the predicted normalization. See Sec. 1.7 for more details.

dataset. Coincidentally, this is is identical to the sensitivity reported in [201]. Sensitivity considerations aside, both measurements are complementary, because, as just discussed, they have a different dependence on systematic uncertainties and model assumptions. A comparison to the result obtained with Medium Energy Starting Events [164] is non-trivial because of the lack of information about the experimental sensitivity.

6.6 Beyond the Single Powerlaw

In this chapter we study possible alternative, more complex astrophysical neutrino flux parametrizations (c.f. Sec. 5.1) and compare them to the baseline single-powerlaw model. Fig. 6.22 shows the best-fit spectra according to the single powerlaw model (top left), the single powerlaw with exponential cutoff (top right), the log-parabolic powerlaw (bottom left) and the broken powerlaw (bottom right) for all cascade-like events. We do not separately show the spectrum obtained assuming a 2-component powerlaw, since, as we will show in Sec. 6.6.4, its best-fit is identical with the single-powerlaw. The

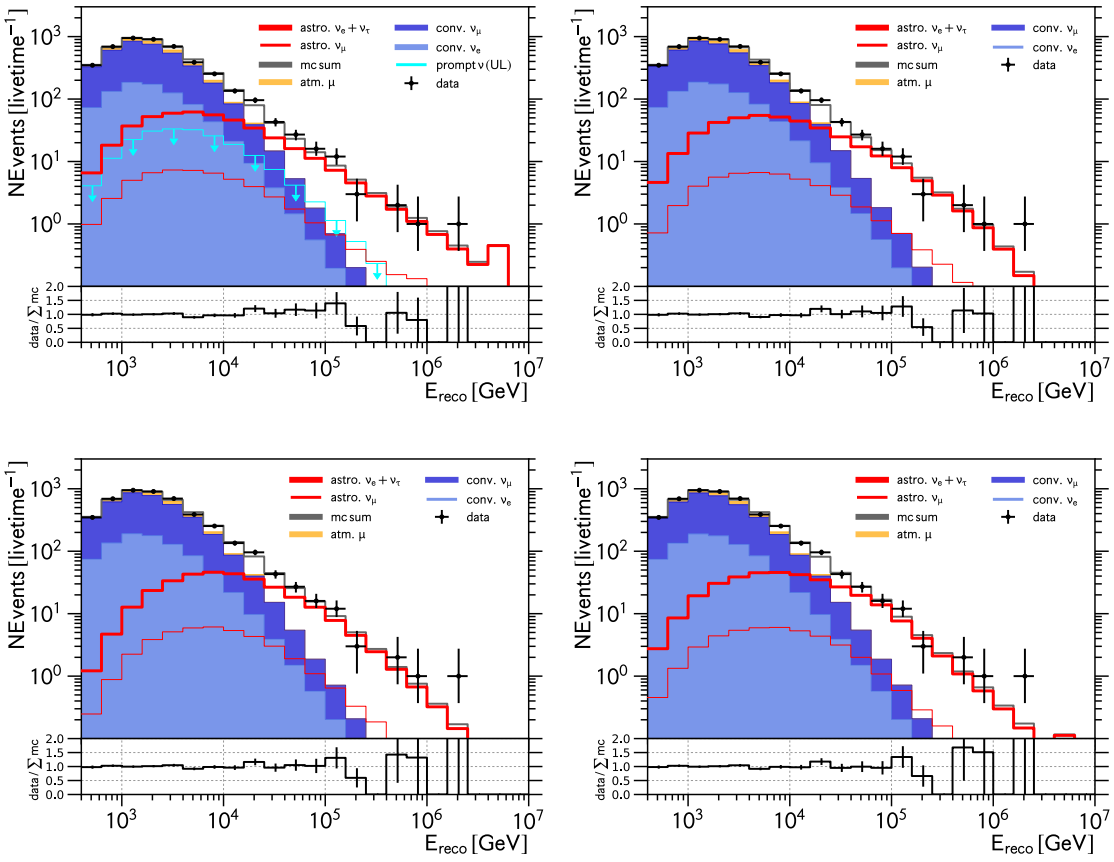


FIGURE 6.22: Comparison between best-fit reconstructed energy spectra according to different astrophysical flux models: single powerlaw (top left), single powerlaw with exp. cutoff (top right), log-parabolic powerlaw (bottom left) and broken powerlaw (bottom right). Cascade events with all directions.

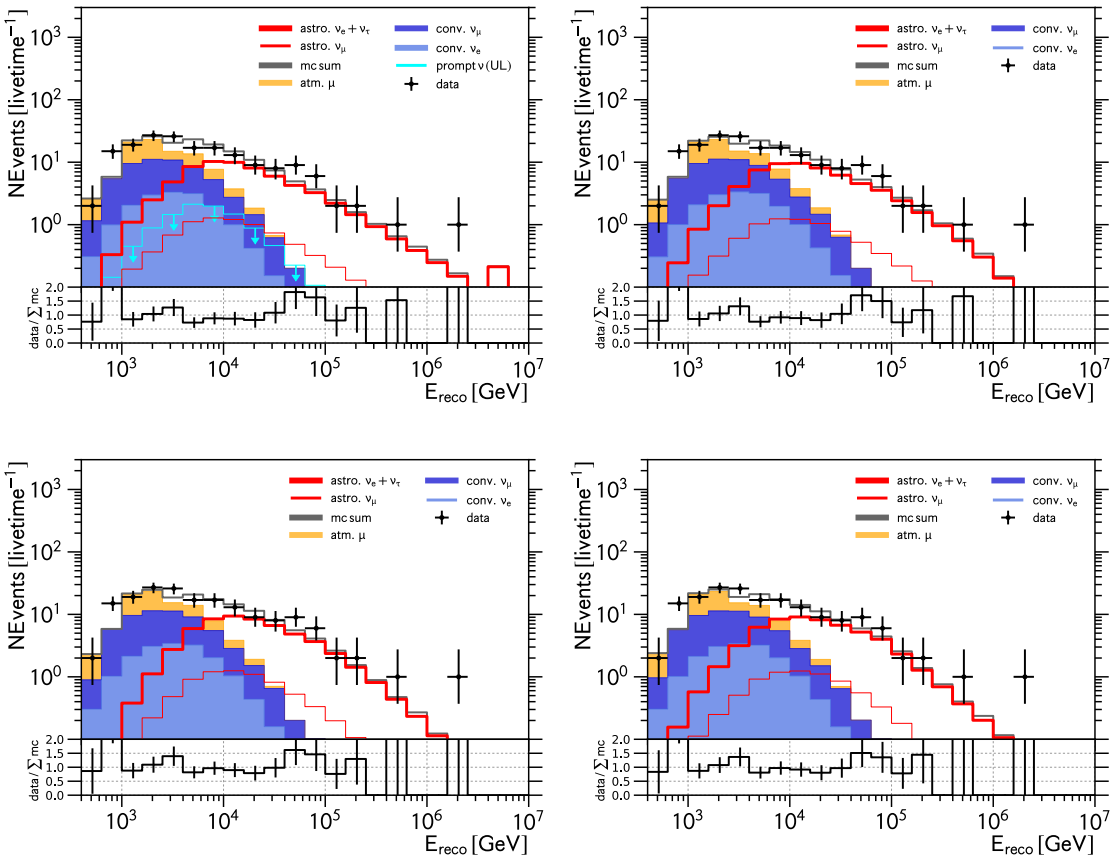


FIGURE 6.23: Comparison between best-fit reconstructed energy spectra according to different astrophysical flux models: single powerlaw (top left), single powerlaw with exp. cutoff (top right), log-parabolic powerlaw (bottom left) and broken powerlaw (bottom right). Cascade events with near vertical directions $0.6 < \cos\theta_{rec}$.

three best-fit alternative models all share one qualitative, general trend: compared to the single-powerlaw model they predict less astrophysical neutrinos at lower energies, more astrophysical neutrinos at intermediate energies (several tens of TeV) and finally less astrophysical neutrinos at highest energies (few hundreds of TeV). These shape differences are most noticeable for near vertically down-going trajectories due to the absence of earth absorption effects at high energies and less atmospheric backgrounds at lower energies. The corresponding spectra, in the same order as before, are shown in Fig. 6.23, while Figs. 6.24 and 6.25 show them for more horizontally, down-going events and up-going events, respectively. We will discuss the results for each model separately and in more detail in their respective sections: Sec. 6.6.1, (single powerlaw with exponential cutoff), Sec. 6.6.2 (log-parabolic powerlaw), Sec 6.6.3 (broken powerlaw) and Sec 6.6.4 (2-component powerlaw). In particular we will study to what extent this dataset does (or does not) prefer each of the more complex models over the single powerlaw baseline assumption by means of separate likelihood ratio tests. Methodically this is similar to testing for an astrophysical single powerlaw flux over the atmospheric background only hypothesis, as discussed in Sec. 6.3. The result is summarized in Tab. 6.5. The smallest p-value ($p = 0.05$) and thus the highest significance (1.6σ) is found for the log-parabolic powerlaw model.

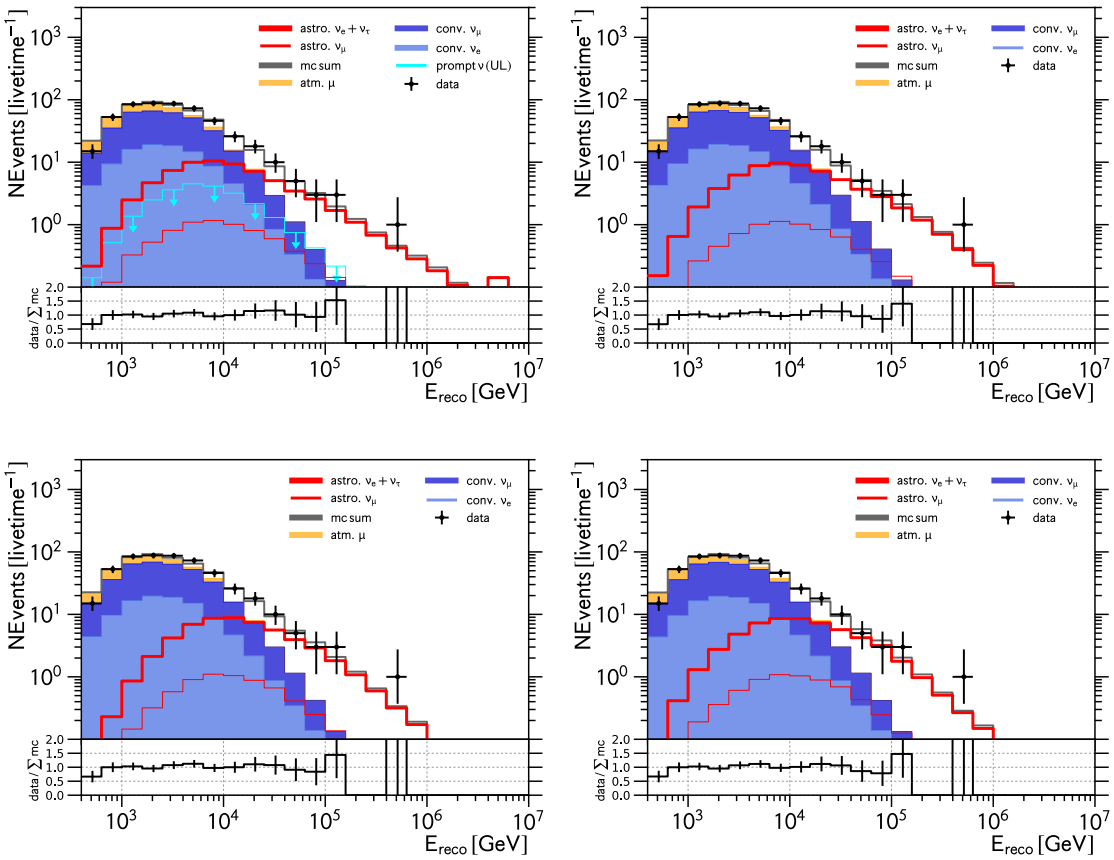


FIGURE 6.24: Comparison between best-fit reconstructed energy spectra according to different astrophysical flux models: single powerlaw (top left), single powerlaw with exp. cutoff (top right), log-parabolic powerlaw (bottom left) and broken powerlaw (bottom right). Cascade events with down-going directions $0.2 < \cos\theta_{rec} \leq 0.6$.

The broken powerlaw ($-2\log\Lambda = 4.6$) fits the data slightly better than the log-parabolic powerlaw ($-2\log\Lambda = 3.1$) when both are compared to the single powerlaw. However due its larger flexibility (one additional parameter) we obtain a smaller significance for the broken powerlaw. While all models describe the data well, none of models provides a significantly better description of the data than the single powerlaw. Thus in this work we did not find any evidence for additional complexity in the astrophysical neutrino flux beyond the single powerlaw.

Finally, it is worth mentioning that the significances given in Tab. 6.5 do not account for the multiple comparisons problem, i.e. the fact that we have tested more than just one model. A simple Bonferroni-adjustment of the p-values is sufficient to control the total false discovery rate [217]. One would either raise the discovery threshold or equivalently multiply each individual p-value by the number of tested hypotheses. For example the adjusted p-value for the log-parabolic powerlaw would read $p_{corr} = 0.2$ (0.8σ).

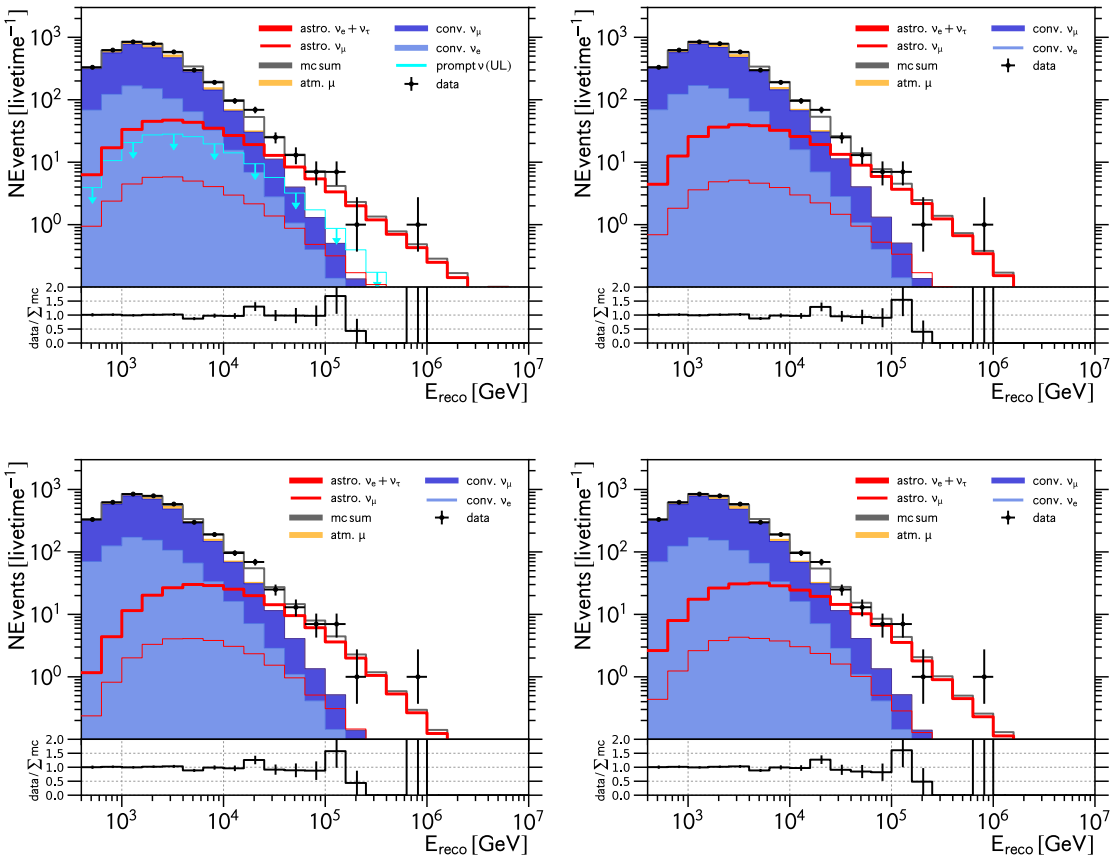


FIGURE 6.25: Comparison between best-fit reconstructed energy spectra according to different astrophysical flux models: single powerlaw (top left), single powerlaw with exp. cutoff (top right), log-parabolic powerlaw (bottom left) and broken powerlaw (bottom right). Cascade events from northern sky $\cos\theta_{rec} \leq 0.2$.

Flux Model (ν_{astro})	Parameters	$-2\log\Lambda$	p-value	significance [σ]
single powerlaw	2	-	-	-
+ exp. cutoff	3	2.0	0.13	1.1
+ log-parabolic shape	3	3.1	0.05	1.6
+ spectral break	4	4.6	0.11	1.2
+ extra component	4	0.0	1.00	0.0

TABLE 6.5: Significance of alternative, mode complex astrophysical flux models over single power-law model as determined from toy experiments (see text).

6.6.1 The Exponential Cutoff

The best-fit values for the parameters of the single powerlaw astrophysical neutrino flux with exponential high energy cutoff are given in Tab. 6.6. We find a cutoff-energy of $\log_{10}(E_{cut}/\text{GeV}) = 6.14^{+0.66}_{-0.40}$, i.e. 1.4 PeV and large uncertainties. Compared to the single, unbroken powerlaw the spectral index hardens to $\gamma = 2.42^{+0.11}_{-0.14}$ and we find an increased

Parameter		Calibration	Result $\pm 1\sigma$
cutoff	$\mathbf{log_{10}(E_{cut}/GeV)}$	-	$(\mathbf{6.14^{+0.66}_{-0.40}})$
spectral index	γ	-	$\mathbf{2.42^{+0.11}_{-0.14}}$
norm astro	ϕ	-	$(\mathbf{1.74^{+0.32}_{-0.28}})$ c.u.
norm conv	ϕ_{conv}	-	$(1.03^{+0.13}_{-0.11}) \times \Phi_{HKKMS06}$
norm prompt	ϕ_{prompt}	-	$< 2.57 \times \Phi_{BERSS}$
norm muon	ϕ_{muon}	-	1.42 ± 0.04
cosmic ray index dev.	$\Delta\gamma_{CR}$	0.00 ± 0.05	0.04 ± 0.03
scattering scale (bulk)	ϵ_{scat}	$1.00 \pm 0.07(^*)$	1.03 ± 0.03
absorption scale (bulk)	ϵ_{abs}	$1.00 \pm 0.07(^*)$	0.99 ± 0.04
scattering scale (HI)	ϵ_{abs}	1.00 ± 0.67	1.68 ± 0.18
DOM efficiency	ϵ_{eff}	0.99 ± 0.10	$1.04^{+0.08}_{-0.07}$
$-2\log\Lambda_{sat}$	62.98		

TABLE 6.6: Best-fit values for assuming the astrophysical neutrino flux to follow a single-powerlaw with exponential cutoff. (*) nuisance parameters are anti-correlated with $\rho = -0.1$

normalization at 100 TeV neutrino energy of $\phi = 1.74^{+0.32}_{-0.28}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$. The increased uncertainties in each parameter are due to correlations with the cutoff-energy. Background-flux and detector-related nuisance parameters remain essentially unchanged. Fig. 6.26 (top) shows the two-dimensional confidence contours for the location of the high energy cutoff in the spectrum and the remaining two parameters: spectral index γ (black) and normalization ϕ (red). As the cutoff-energy decreases, the data prefers harder spectral indices. The two-dimensional confidence contour for the powerlaw parameters (γ and ϕ) is shown in Fig. 6.26 (bottom left) and appears significantly larger than the one corresponding to the unbroken, single powerlaw from before (dashed, black line, c.f. Fig. 6.6). This has interesting implications. It is very obvious that the possibility of a high energy cutoff increases uncertainties in possible flux predictions at high energies. It might, however, be considered less obvious by some, that this also increases uncertainties in flux predictions at lowest energies - as we just demonstrated in Fig. 6.26.
We compared this model to the single, unbroken powerlaw with the following likelihood ratio test:

$$\begin{aligned} H_0 &: \text{powerlaw w/o cutoff } (''\log_{10}(E_{cut}/GeV) = \infty'') \\ \text{against} \quad H_1 &: \text{powerlaw w/ cutoff } (''\log_{10}(E_{cut}/GeV) < \infty'') \end{aligned} \tag{6.15}$$

The corresponding distribution of the test-statistic $-2\log\Lambda$ is shown in Fig. 6.26 (bottom right) and has been determined from simulating replicated toy-experiments assuming the best-fit values from Tab. 6.1. In 14% of the cases we find a small preference for the existence of a high energy cutoff that is equal or larger than the one observed ($-2\log\Lambda \geq 2.0$). The shape of the test-statistic does not appear to follow the distribution that one might naively have expected (at least for large sample sizes). Technically, the two models are not nested, since the alternative model does not reduce to the null model within its parameter space (infinity is not a real number). Ignoring technicalities, one might extend the real line to include $+\infty$ as a boundary point and potentially consider a theorem by Chernoff [193]. In this case one would expect the sampling distribution to approach eq.

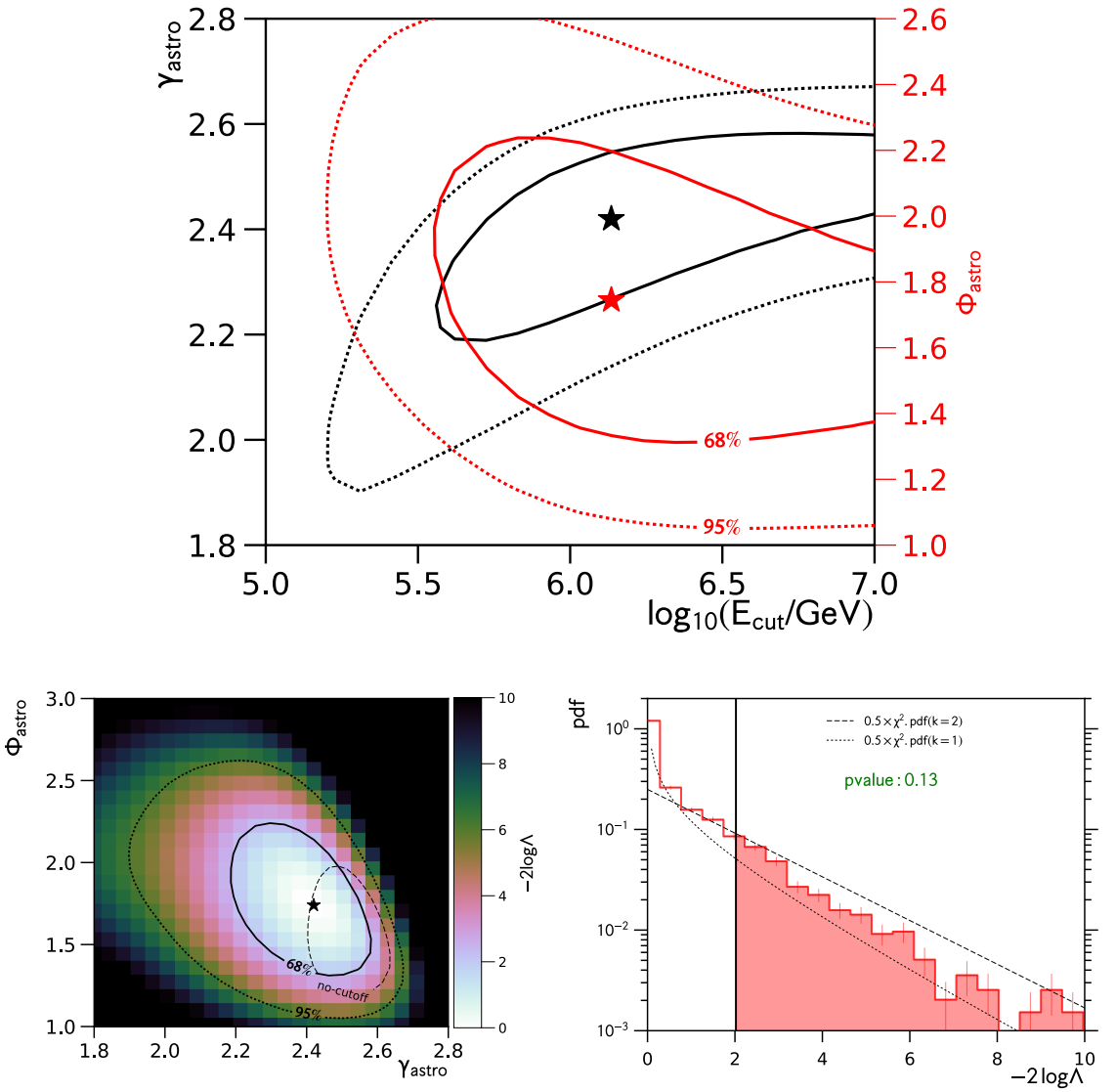


FIGURE 6.26: Top: confidence contours for cutoff energy and normalization in red (spectral index in black). Bottom left: confidence contour for normalization and spectral index. Bottom right: test-statistic distribution of likelihood ratio to calculate significance w.r.t to single power-law without cutoff.

(6.4) for large sample sizes. While our simulated test-statistic distribution does show signs of boundary behavior, i.e. it is strongly peaked at/near $-2\log\Lambda = 0$, its tail behavior deviates from eq. (6.4). One should keep in mind, however, that information about the cutoff parameter is extracted at highest energies where data data is sparse and, in particular, asymptotic considerations do not apply.

Finally it is worth having a closer look at the impact of the Glashow Resonance (c.f. Sec. 1.8) at $E_\nu = 6.3\,\mathrm{PeV}$ and the non-observation thereof on the marginal preference of the spectral cutoff discussed above. This can be illuminated by replacing the exponential cutoff by an instantaneous cutoff beyond which the flux vanishes ($\Phi(E) = 0$, $E > E_{cut}$). Fig. 6.27 (black line) shows the profile-likelihood function for the (instantaneous) cutoff

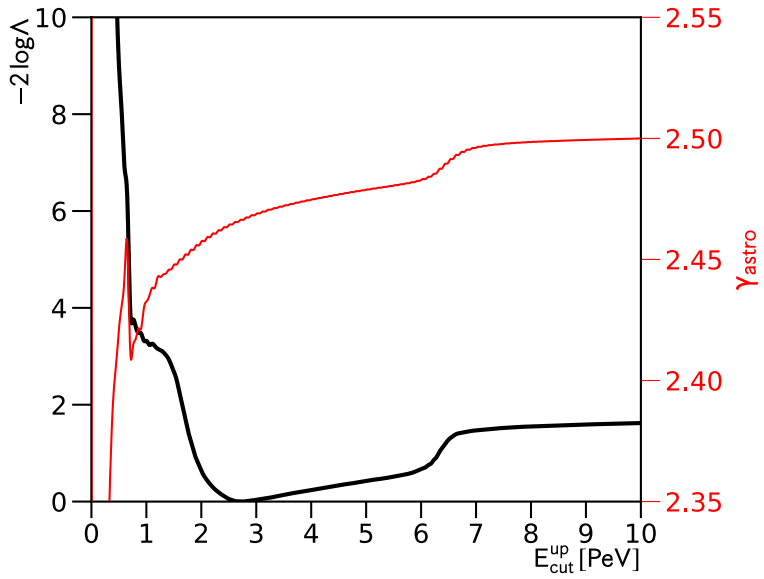


FIGURE 6.27: Likelihood function for lternative high energy cutoff model: astrophysical flux vanishes instantaneously for $E > E_{cut}$. Detector-related systematics not included.

energy (no detector systematics included). The cutoff energy increased compared to the previous estimate based on the exponential cutoff, since the flux has to be non-zero at energies where we have observed corresponding events. The influence of the Glashow resonance is visible as a steepening of the profile likelihood function at $\sim 6.3\,\mathrm{PeV}$. At even higher energies it approaches a constant, since our dataset has no sensitivity to the low fluxes expected from a single powerlaw at these very high energies. In this simplified cutoff model, roughly half of the (insignificant) preference (half of the test-statistic value at the plateau) for a cutoff is contributed by the non-observation of the Glashow Resonance. Should future IceCube data with larger statistic provide stronger evidence for a cutoff based on the non-observation of the Glashow Resonance, it will be necessary to include uncertainties on the astrophysical $\nu_e : \bar{\nu}_e$-ratio. If the spectrum is depleted from electron anti-neutrinos the Glashow Resonance signal will be suppressed. Finally, for the interested reader, we also show the powerlaw spectral index as function of instantaneous cutoff energy (red line). The impact of the non-observation of the Glashow Resonance on the spectral index is negligible ($\Delta\gamma < 0.02$).

6.6.2 The Log-Parabola

The best-fit parameters for the log-parabolic powerlaw (eq. (5.8)) can be found in Tab. 6.7. We find a non-zero value for the parameter that controls the log-parabolic curvature $b = 0.096^{+0.077}_{-0.060}$. Furthermore we obtain a spectral index $\gamma_{astro} = 2.62 \pm 0.12$ and a flux normalization of $\phi_{astro} = (1.77^{+0.33}_{-0.29})\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$, both measured/defined at a neutrino energy of 100 TeV. Fig. 6.28 (top) shows the two-dimensional confidence contours for the curvature parameter b and either one of the remaining two powerlaw parameters, spectral index γ_{astro} (black) and normalization ϕ_{astro} (red). Both show a weak correlation with the curvature parameter. The inclusion of the additional parameter

Parameter		Calibration	Result $\pm 1\sigma$
log-slope	**b**	-	$(\mathbf{0.096^{+0.077}_{-0.060}})$
spectral index	γ	-	$\mathbf{2.62 \pm 0.12}$
norm astro	ϕ	-	$(\mathbf{1.77^{+0.33}_{-0.29}})$ c.u.
norm conv	ϕ_{conv}	-	$(1.04^{+0.13}_{-0.11}) \times \Phi_{HKKMS06}$
norm prompt	ϕ_{prompt}	-	$< 2.97 \times \Phi_{BERSS}$
norm muon	ϕ_{muon}	-	1.42 ± 0.04
cosmic ray index dev.	$\Delta\gamma_{CR}$	0.00 ± 0.05	0.04 ± 0.03
scattering scale (bulk)	ϵ_{scat}	1.00 ± 0.07(*)	1.03 ± 0.03
absorption scale (bulk)	ϵ_{abs}	1.00 ± 0.07(*)	0.98 ± 0.04
scattering scale (HI)	ϵ_{abs}	1.00 ± 0.67	1.67 ± 0.18
DOM efficiency	ϵ_{eff}	0.99 ± 0.10	$1.04^{+0.08}_{-0.07}$
$-2 \log \Lambda_{sat}$	61.87		

TABLE 6.7: Best-fit values for assuming the astrophysical neutrino flux to follow a log-parabolic power-law. (*) nuisance parameters are anti-correlated with $\rho = -0.1$

again increases uncertainty about the spectral index and normalization as shown in Fig. 6.28 (bottom left).
We again perform a likelihood ratio test to compare this model with the baseline single powerlaw assumption:

$$H_0 : b = 0 \quad \text{against} \quad H_1 : b > 0 \tag{6.16}$$

This test satisfies the conditions for a theorem by Chernoff [193], and thus the test-statistic distribution is expected to approach eq. (6.4) as the size of the sample becomes large. We again derived the test-statistic distribution for sample sizes comparable to ours (within Poisson fluctuations) from simulated toy-experiments assuming the true values to be given by Tab. 6.1. The resulting distribution (4200 toy-datasets) is shown in Fig. 6.28 (bottom right). The distribution appears very similar but not identical to the asymptotic expectation from eq. (6.4). This is probably again due the fact that the parameter b is sensitive to the high energy region and correspondingly the small number of events there. Finally, we obtain a p-value of $p = 0.049$.

6.6.3 The Broken Powerlaw

The best-fit parameters for the broken powerlaw (eq. (5.6)) can be found in Tab. 6.8. We find a break energy of $\log_{10}(E_b/GeV) = 4.85^{+0.33}_{-0.40}$, i.e. a spectral break at 71 TeV. Compared to the single powerlaw we find a harder best-fit spectral index of $\gamma_1 = 2.27^{+0.18}_{-0.34}$ at energies below the break energy and a softer best-fit spectral index of $\gamma_2 = 2.92^{+0.41}_{-0.25}$ at energies above. The latter well-reproduces the soft spectral index measured with High Energy Starting Events [204] above 60 TeV (c.f. discussion in Sec. 6.2). We further find a flux normalization at a neutrino energy of 100 TeV of $\Phi = 1.86^{+0.55}_{-0.42}\,\text{GeV}^{-1}\text{cm}^{-2}\text{s}^{-1}\text{sr}^{-1}$. Fig. 6.29 (top) shows the two-dimensional profile likelihood contours for the break energy ($\log_{10}(E_b)$) and either of the two spectral indices (γ_1, γ_2). The uncertainties are very large, especially as the break energy goes to small (large) energies. In those regions of the parameter-space it is difficult to constrain the spectral index of the lower

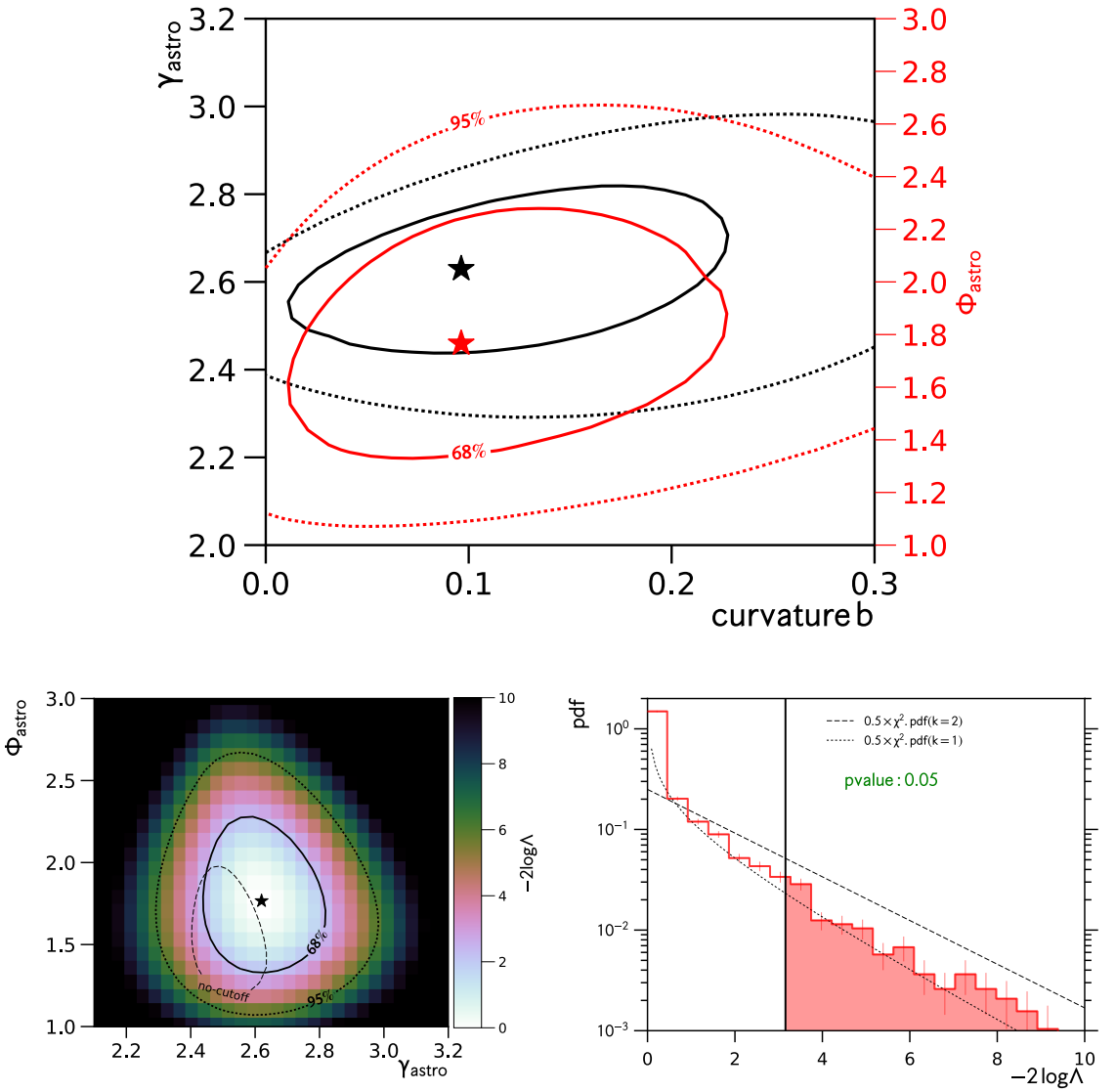


FIGURE 6.28: Top: confidence contours for log-parabolic curvature and normalization in red (spectral index in black). Bottom left: confidence contour for normalization and spectral index. Bottom right: test-statistic distribution of likelihood ratio to calculate significance w.r.t to single power-law without cutoff.

(higher) energies, γ_1 (γ_2). The degenerate configuration that recovers the single power-law $\gamma_1 = \gamma_2 = 2.53$ is visualized by the black, dashed line. If both spectral indices lie on this line, the break energy is not defined and has no meaning. This is similar to case of the single powerlaw at $\Phi = 0$, where the spectral index becomes undefined. We also calculated a profile-likelihood contour for the two spectral indices (γ_1, γ_2) to facilitate a better visual comparison to the degenerate point. However, this particular calculation suffered from instabilities due to multiple minima along the break energy dimension, leading to confusion during numerical minimization. To avoid these numerical problems, we performed a similar calculation using a Bayesian approach and ensemble MCMC.

Parameter		Calibration	Result $\pm 1\sigma$
break energy	$\mathbf{\log_{10}(E_b/GeV)}$	-	$(\mathbf{4.85^{+0.33}_{-0.40}})$
spectral index (1)	$\boldsymbol{\gamma_1}$	-	$\mathbf{2.27^{+0.18}_{-0.34}}$
spectral index (2)	$\boldsymbol{\gamma_2}$	-	$\mathbf{2.92^{+0.41}_{-0.25}}$
norm astro	$\boldsymbol{\phi}$	-	$(\mathbf{1.86^{+0.55}_{-0.42}})$ c.u.
norm conv	ϕ_{conv}	-	$(1.03^{+0.13}_{-0.11}) \times \Phi_{HKKMS06}$
norm prompt	ϕ_{prompt}	-	$< 2.57 \times \Phi_{BERSS}$
norm muon	ϕ_{muon}	-	1.42 ± 0.04
cosmic ray index dev.	$\Delta\gamma_{CR}$	0.00 ± 0.05	0.04 ± 0.03
scattering scale (bulk)	ϵ_{scat}	1.00 ± 0.07(*)	1.03 ± 0.03
absorption scale (bulk)	ϵ_{abs}	1.00 ± 0.07(*)	0.98 ± 0.04
scattering scale (HI)	ϵ_{abs}	1.00 ± 0.67	1.67 ± 0.18
DOM efficiency	ϵ_{eff}	0.99 ± 0.10	$1.04^{+0.08}_{-0.07}$
$-2\log\Lambda_{sat}$	60.44		

TABLE 6.8: Best-fit values for assuming the astrophysical neutrino flux to follow a broken power-law. (*) nuisance parameters are anti-correlated with $\rho = -0.1$

Assuming uniform prior distributions for the unknown astrophysical parameters and ignoring detector-related systematic uncertainties to reduce the number of dimensions, we marginalized over all remaining nuisance parameters and obtain the joint two-dimensional posterior distribution for γ_1 and γ_2 shown in Fig. 6.29 (bottom right). We find the uncertainties derived in this way to be slightly larger than their frequentist counterparts. In particular the 68% highest posterior density contour touches the degeneracy axis (white dashed line).
To estimate the significance corresponding to the improved fit of this model over the single powerlaw we again employ a likelihood ratio test:

$$H_0 : \gamma_1 = \gamma_2 \quad \text{against} \quad H_1 : \gamma_1 \neq \gamma_2 \tag{6.17}$$

This test does not involve any parameter space boundaries and we therefore do not expect any related effects to influence the test-statistic ($-2\log\Lambda$) distribution. Assuming a fixed value for the break energy we expect from Wilks' theorem the distribution to be χ^2_1. This expectation is an excellent match to our simulation based estimate ($\log_{10}(E_b/GeV) \in \{4.3, 4.6, 4.9\}$ fixed), see Fig. 6.30 (top left). However, for the realistic case of the break energy being unknown, the distribution is unknown. In this case Wilk's theorem does not apply because, as discussed, the break energy is undefined under H_0. Fig. 6.30 (top right) shows the distribution of $-2\log\Lambda$ obtained from simulating replicated toy-datasets and perform the hypothesis test with variable break energy. We determine a p-value of 0.10. Since this model is complex and we previously observed multiple minima along the break energy dimension, we did not fully trust this result. As discussed in the context of rejecting the atmospheric background only hypothesis (c.f. Sec. 6.3) one can (under suitable conditions) relate the test involving a variable parameter, undefined under H_0, to a simpler problem, in which this parameter has a known, constant value. However a *look elsewhere correction* is required [213][214][211]. The p-value can be determined from

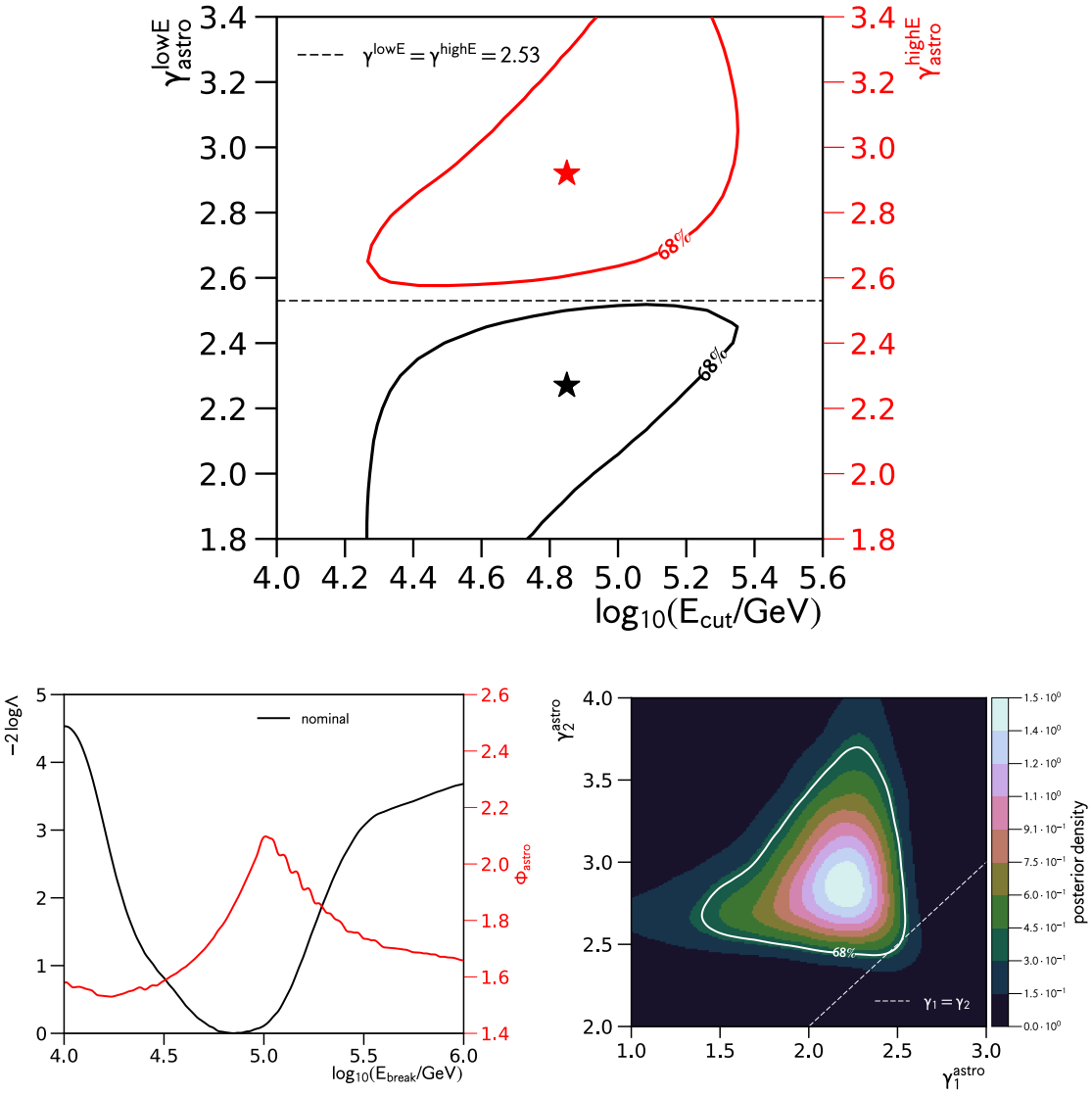


FIGURE 6.29: Top: confidence contours for break energy and high energy spectral index in red (low energy spectral index in black). Bottom left: profile likelihood function and flux normalization as function of break energy. Bottom right: Bayesian credible region for both spectral indices (low and high energy).

[212]

$$p = P\left(\chi_1^2 > \xi_{obs}\right) + E[U(c_0)|H_0] \times \exp\left(-\frac{\xi_{obs} - c_0}{2}\right) \tag{6.18}$$

where $E[U(c_0)|H_0]$ is again the expected number of up-crossings of the in this particular case corresponding χ_1^2-random process[9]. Defining a search region of $\log_{10}(E_b/GeV) \in [4.0, 6.0]$ and choosing $c_0 = 0.5$ we find $E[U(c_0)|H_0] = 0.60$, as shown in Fig. 6.30

[9]Eq. (6.5) appears slightly different, because of the parameter boundary at $\Phi = 0$ that yields a $\frac{1}{2}\chi_1^2 + \delta(0)$-random process (for the test discussed in Sec. 6.3) instead of a χ_1^2-random process (for the test discussed here).

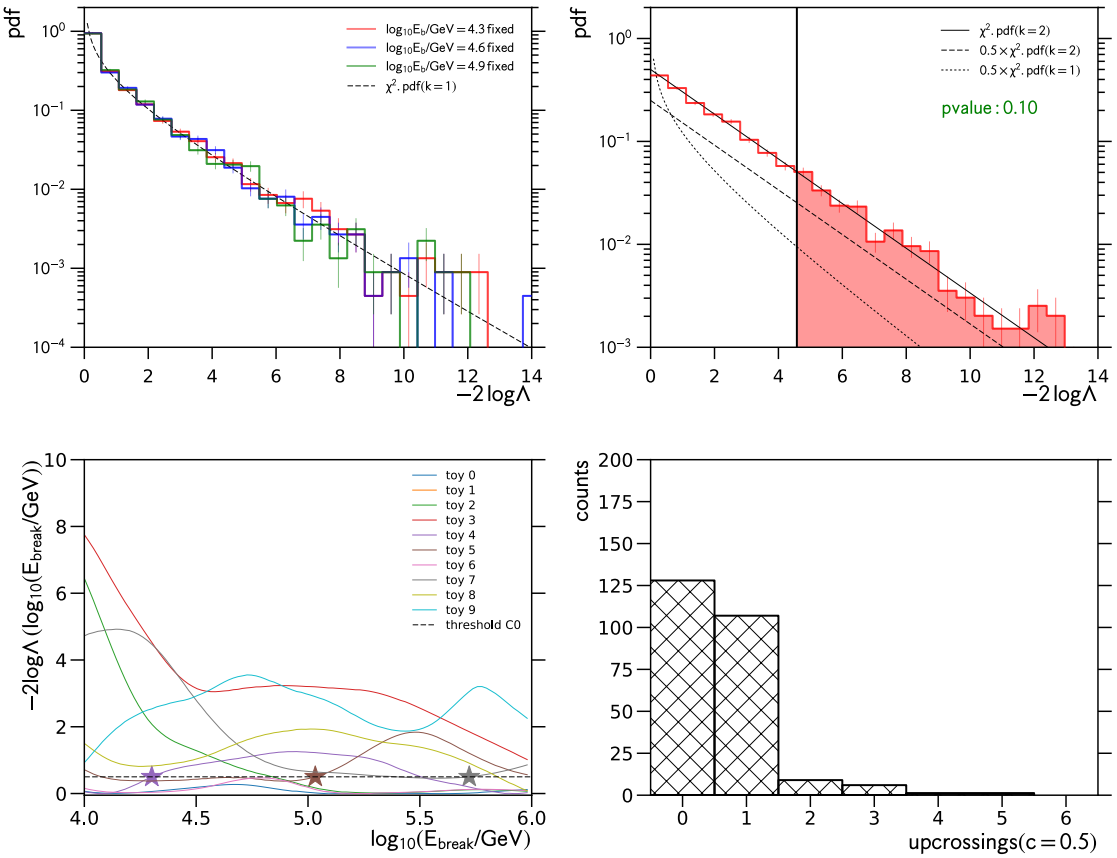


FIGURE 6.30: Significance calculation for broken powerlaw model. Top left: test-statistic distribution for fixed break energy. Top right: test-statistic distribution for variable break energy. Bottom left: up-crossing of LRT random process. Bottom right: distribution of number of up-crossings.

(bottom left: random process realizations, bottom right: distribution of up-crossings). From eq. (6.18) we find

$$p = 0.11 \tag{6.19}$$

which compares very well to the previous estimate and is the final result of this test. For completeness we determine a trials factor of 3.4 for the search window (break energy) mentioned above.

6.6.4 The 2-Component Powerlaw

The 2-component powerlaw, eq. (5.5), does not provide a better fit to this dataset than the single powerlaw, despite introducing two additional parameters. The best-fit value for the normalization of the additional, harder flux component is $\Phi_{2,0} = 0\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$ and thus no high energy spectral hardening is observed. The model reduces to and is identical with the single powerlaw. The best-fit parameters can be found in Tab. 6.1. Since tow

component models have received considerable attention in the literature [218][219][220], we will further examine this null result in the following. Fig. 6.31 (top) shows the two-dimensional profile likelihood confidence contours for the two new parameters in eq. (5.5) (mixture fraction α and spectral hardening $\Delta\gamma$). At the boundaries ($\alpha \to \{0, 1\}$ or $\Delta\gamma = 0$) the model is degenerate and reduces to the single powerlaw. For small (large) values of the mixture fraction this dataset can not meaningfully constrain the amount of spectral hardening $\Delta\gamma$, because the hard (soft) component is subdominant compared to the "background" from the soft (hard) component. We can exclude 2-components powerlaws that for each component predict roughly equal intensities at 100 TeV ($0.3 < \alpha < 0.7$)

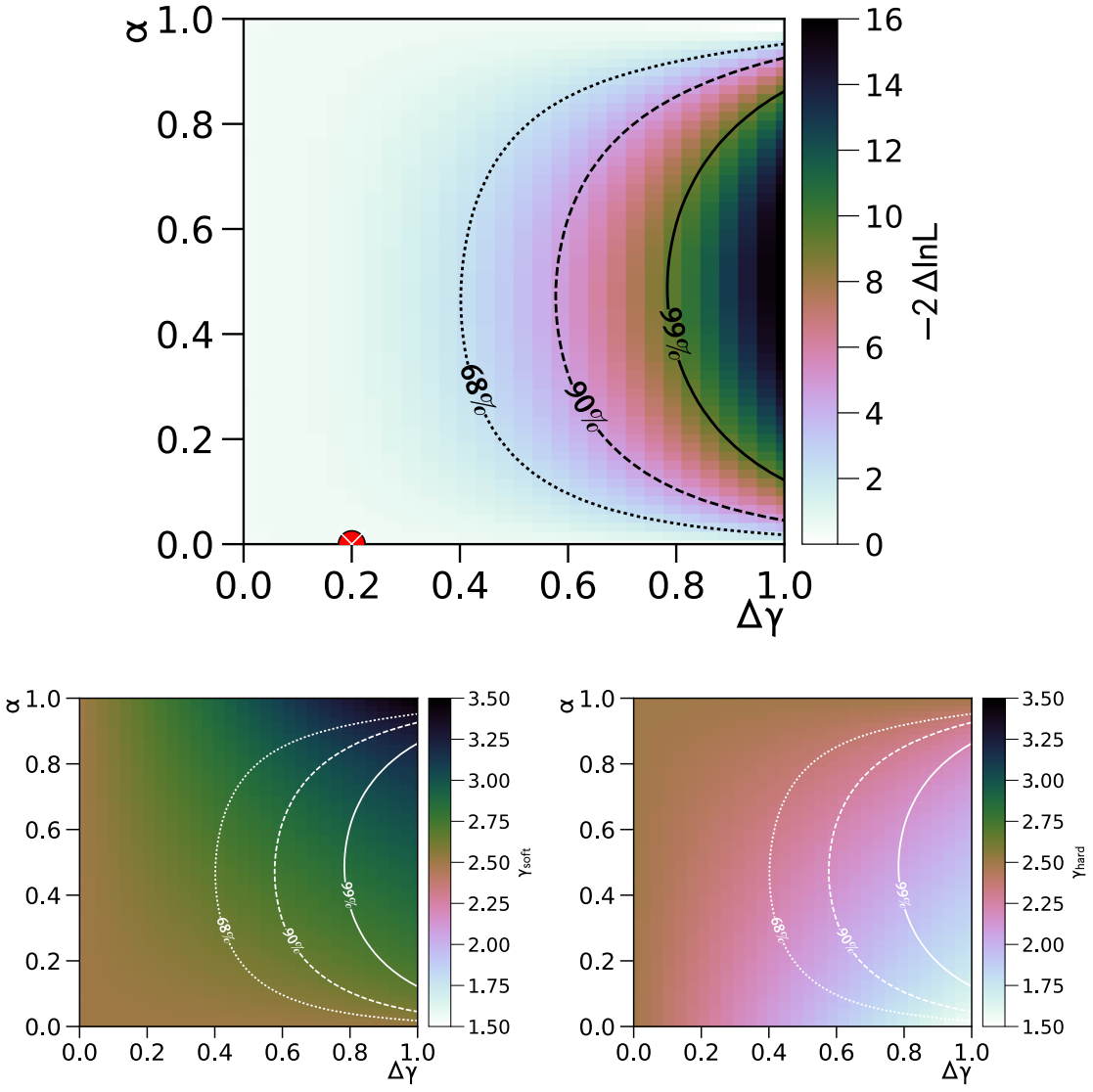


FIGURE 6.31: Uncertainties related to the two new parameters in the 2-component parameter model: mixture fraction α and spectral hardening $\Delta\gamma$. Top: profile likelihood function. The best-fit value is marked by the red point. Bottom: soft (left) and hard (right) spectral indices as function of α and $\Delta\gamma$. Detector-related systematics not included.

and spectral hardening $\Delta\gamma$ greater than ~ 0.6 at 90% C.L.. Fig. 6.31 (bottom) shows the corresponding conditional best-fit values for the soft (left) and hard (right) spectral indices. If $\Delta\gamma$ becomes large, i.e. $\alpha \to \{0, 1\}$ the spectral index of the dominant component approaches best-fit value from the single powerlaw (~ 2.5), while the subdominant component realizes very large (soft component) or very small values (hard component), thus generating the large spectral difference between both components. We should mention that we did not include detector-related systematic uncertainties in the calculation of these contours to speed up the numerical minimizations. We assumed a nominal detector response except for the non-standard hole-ice scattering length of 30 cm.
In order to facilitate a better comparison to the latest measurement using muon neutrinos from the northern hemisphere [205], we would like to change the model parametrization from eq. (5.5) (useful for hypothesis testing purposes) to the one of eq. (5.3). For numerical reasons, this is easier done within a MCMC based calculation. Hence we performed the calculation using a Bayesian approach. We employ prior probability distributions (including the detector-related systematic parameters) that reflect information obtained in auxiliary calibration measurements (c.f. Sec. 5.3 and 5.8). For the truly unknown parameters ($\Phi_{prompt}, \Phi_{astro}, \gamma_{soft}, \alpha, \Delta\gamma$) we assume uniform prior distributions[10]. The resulting 68% highest marginal posterior density regions for both components, soft (red) and hard (blue), are shown in Fig. 6.32 (top left). Also shown is the result based on muon neutrinos [205] (green). Superficially the agreement is improved, as the contours (green, blue) appear lass separated compared to the single powerlaw assumption (c.f. Sec. 6.2 and Fig. 6.13), due the increased uncertainties. Fig. 6.32 (top right) visualized the strong anti-correlation between both components. When the hard component realizes its maximum intensity, the flux from the soft component vanishes. In particular the distribution appears bi-modal, where each peak represents a solution that reduces to the single powerlaw.

While no evidence for 2-components was found, it is still interesting to ask what the two-component flux would look like, if information from the muon neutrino measurement was added. While this is best done within a global fit [199][202], we can provide an approximate solution by treating the result from the muon neutrino measurement [205] (green contour) as a prior distribution for the harder of the two flux components (blue). The result is shown in Fig. 6.32 (2nd row). The posterior distribution of the

[10]We did not perform a sensitivity analysis to these choices. If in the future the evidence for this 2-component model becomes more convincing, this would become necessary. In particular we would advice to replace the uniform prior for Φ_{prompt} by one derived from theory, e.g. using [102], in order to a-priori penalize unrealistically large prompt neutrino fluxes. Uniform priors do not represent a realistic state of knowledge, if true effect sizes are expected to be small [221].

Parameter		Result	68%-HPD interval
Norm (1)	Φ_{hard}	1.10 c.u.	[0.89, 1.40]
spectral index (1)	γ_{hard}	2.29	[2.20, 2.38]
Norm (2)	Φ_{soft}	0.56 c.u.	[0.12, 0.75]
spectral index (2)	γ_{soft}	2.72	[2.49, 2.86]
$-2 \log \Lambda_{sat}$	68.12		

TABLE 6.9: Credible intervals for each parameter in the two-component scenario - includes the ν_μ-tracks prior [205].

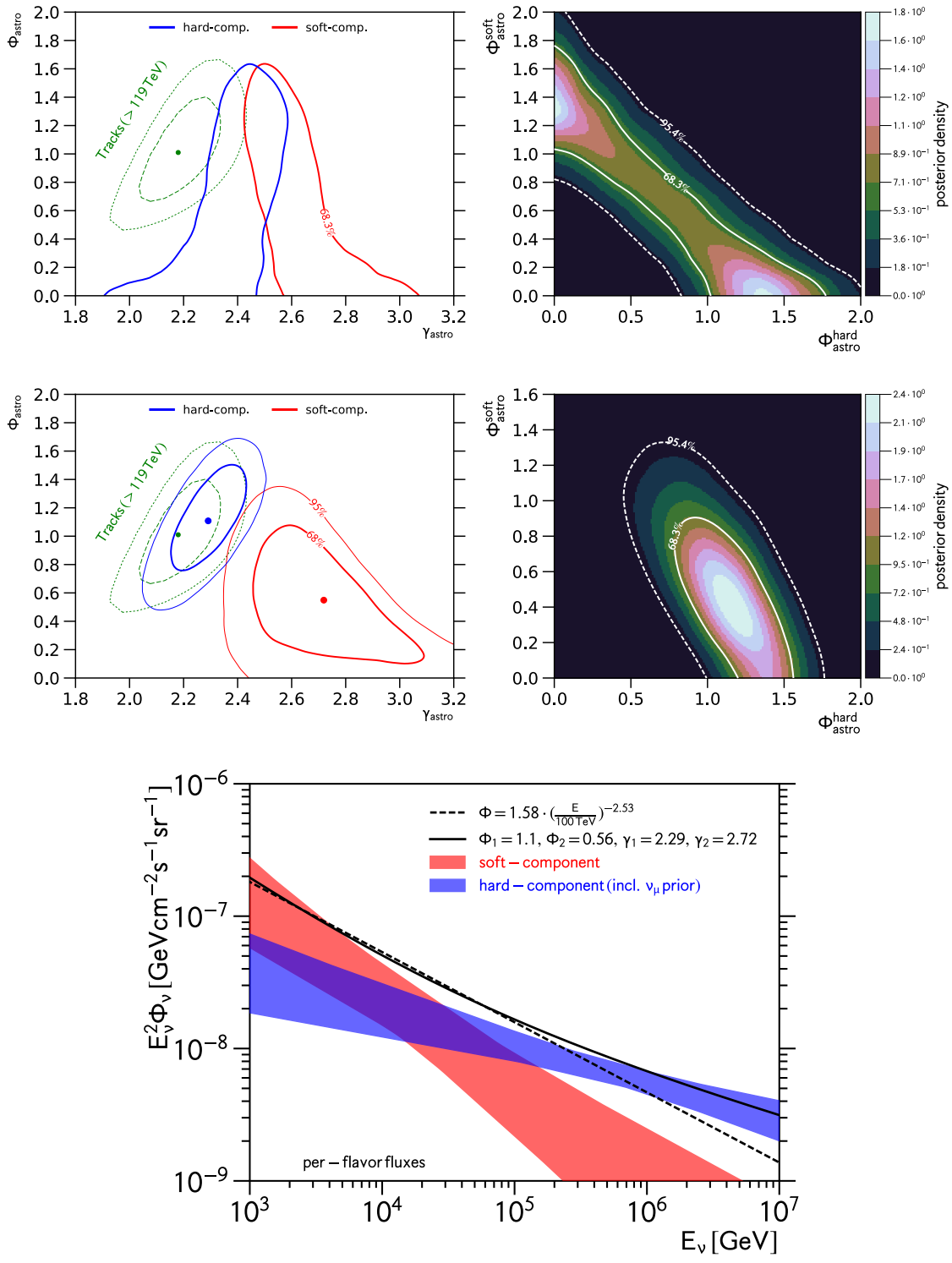


FIGURE 6.32: Highest posterior credible regions for two component powerlaw. Top: uniform prior distributions for $\Phi_{1,2}$ and $\gamma_{1,2}$. Two dimensional contours for each component (left). Anti-correlation between normalizations of both components (right). Middle: same above, but using ν_μ-track measurement [205] (green) as prior for the hard component (blue). Bottom: corresponding flux.

hard component (blue) does not show a strong pull away from its prior distribution (green), as would be expected if the data conflicts with the prior. The maximum likelihood (maximum a posteriori probability) point estimate for the hard component is $\Phi_{hard} = 1.10 \cdot 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$ at 100 TeV and $\gamma_{hard} = 2.29$. In addition we find a subdominant soft component characterized by $\Phi_{soft} = 0.56 \cdot 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$ at 100 TeV and $\gamma_{soft} = 2.72$. Highest posterior density credible intervals for each parameter are given in Tab. 6.9. Obviously, this solution does not describe this dataset as well as the single powerlaw does, since the best-fit 2-component powerlaw reduces to the single powerlaw if the information provided by the muon neutrino measurement is ignored. The difference in fit-quality between the best-fit solution (Tab. 6.1) and the "priored" solution (Tab. 6.9) equals 3.1 units of $-2\log\Lambda$, corresponding to a p-value of ~ 0.5[11]. Thus, assuming that nature indeed realized a 2-component powerlaw, the preference for the single powerlaw configuration would not be significant, and, in particular, the "priored" solution is consistent with this dataset. Fig. 6.32 (bottom) compares the "priored" double powerlaw solution (solid black line) to the flux obtained assuming a single powerlaw (dashed black line). At neutrino energies below 100 TeV, where this analysis most sensitive, both solutions are essentially indistinguishable. At higher energies, the prior from [205] requires a larger flux compared to the single powerlaw. Fig. 6.32 also shows the decomposition of the total double powerlaw flux (solid black line) into its 2-components: soft component (red band) and hard component (blue band).
It remains to be evaluated, however, whether or not the subdominant soft component in this model would conflict with the muon neutrino data at low energies. Finally, we would like to point out that the discussion above can not replace a complete, joint analysis in form of a global fit.

6.6.5 The Piecewise Model

The piecewise-model allows for a less model-dependent analysis of the astrophysical neutrino spectrum. It assumes a (logarithmic) step-function for the astrophysical energy flux in neutrinos ($E_\nu^2\Phi_\nu$) and is given by eq. (5.12). Due to the large number of parameters and thus the large flexibility of this model we have not been able to include detector systematic uncertainties in the fit. We assume a nominal detector response except for an alternative hole-ice scattering length of 30 cm that appears strongly preferred by this analysis. Fig. 6.33 shows the best-fit piecewise model (horizontal black bars) and 68% C.L. confidence intervals (or upper limits) on the normalization parameters for each segment (vertical black bars). For comparison the same figure shows the single-powerlaw best fit from Tab 6.1 (black line) and corresponding point-wise (black dots) as well as joint (red band) 68% C.L. confidence bands including all detector systematic uncertainties. Additionally we highlighted the single powerlaw energy range derived in eq. (6.11) in red. Both models compare well, where the data can meaningfully constrain all relevant parameters (i.e $\sim$ 10 TeV $-$ 150 TeV). At higher energies the available data statistics is scarce and the piecewise model does not provide additional information about the spectral behavior of the flux, except for the total intensity. Similarly, uncertainties are

[11] This number ignores possible (near-)boundary effects, that invalidate Wilks' theorem but are not expected to alter the conclusion.

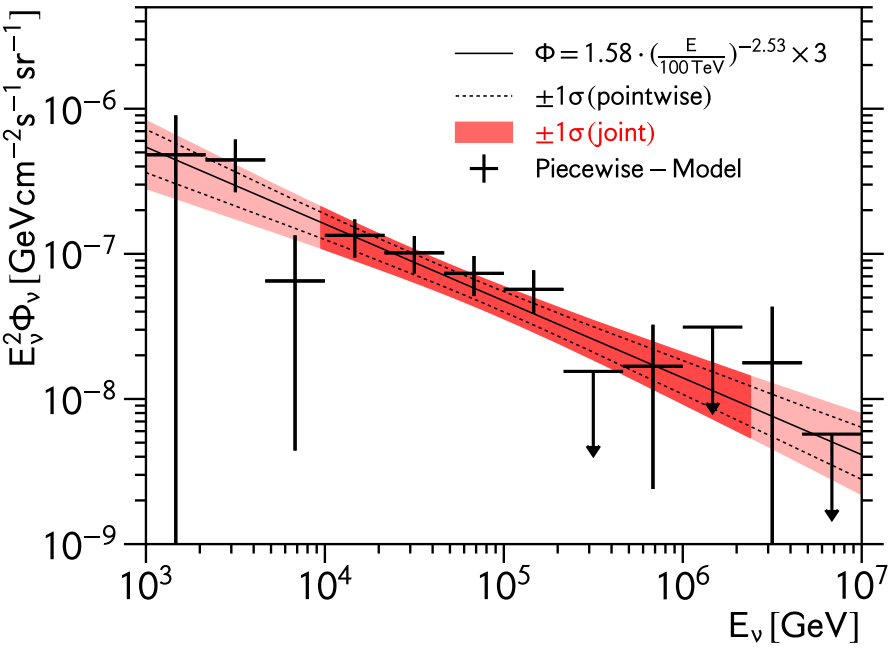


FIGURE 6.33: Best-Fit astrophysical neutrino flux according to the piecewise flux model (black). compared to single powerlaw flux measurement (red band).

large at low energies, primarly because of the dominant conventional atmospheric neutrino background and corresponding fluctuations. Without formal model comparison it appears obvious that the piecewise model fit is not providing evidence against the single powerlaw model because of the good agreement between both.
Fig. 6.34 shows the same comparison for the different spectral models beyond the single powerlaw, discussed in Sec. 6.6. While in agreement with all of them, they highlight one particular aspect: uncertainties on the astrophysical flux increase very rapidly at the low and high energy ends, especially if joint coverage is required - for example if the flux needs to be extrapolated in order to compare to other experiments (c.f. multi-messenger constraints, e.g. Sec. 1.6). Already the log-parabolic flux model (top right), which only adds one more parameter, shows significantly larger uncertainties towards low and high energies than the single powerlaw model.

6.6.6 The 2-Hemisphere Model

All of the previous models left one assumption un-questioned: They all assumed that the astrophysical neutrino flux is isotropic. The 2-hemisphere model, eq. (5.11), allows for testing of that assumption, especially to what extent some of the previous results depend on it. A possible north-south asymmetry in the astrophysical neutrino has been reported by a recent, combined IceCube analysis [199], albeit large uncertainties and low significance (1.1σ). In particular a hard spectrum was observed in the northern sky, $\gamma_N = 2.0^{+0.3}_{-0.4}$, with a small normalization of $\Phi_N = 0.7^{+1.0}_{-0.5} \cdot 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$. The spectral index obtained from the southern hemisphere appeared similar to the one

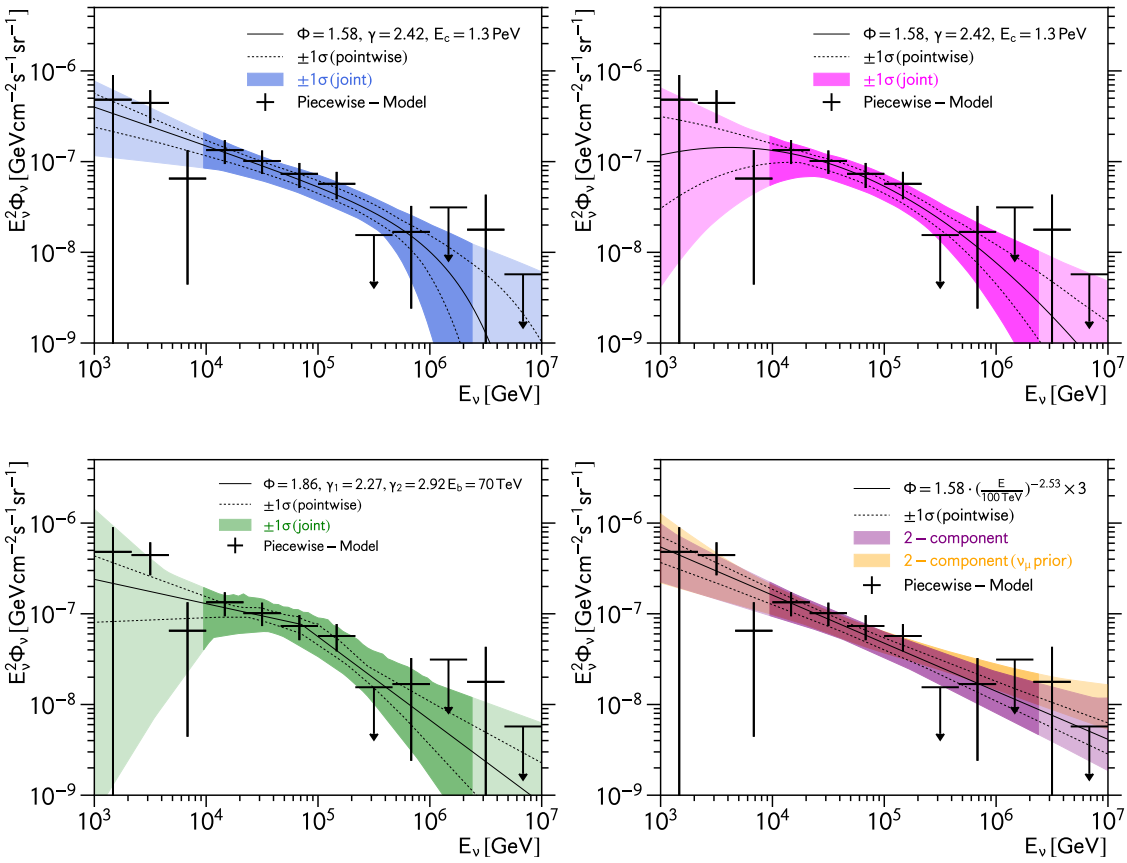


FIGURE 6.34: Best-Fit astrophysical neutrino flux according to the piecewise flux model (black). Comparison to other flux models: single powerlaw w. cutoff (top left, blue), log-parabolic powerlaw (top right, magenta), broken powerlaw (bottom left, green) and two-component powerlaw (bottom right, purple)

obtained in this work $\gamma_S = 2.56 \pm 0.12$ but with larger normalization of $(2.3 \pm 0.5) \cdot 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$. Such asymmetry has subsequently been proposed to explain the difference between the preferred spectra by tracks (muon neutrinos) and cascades (electron and tau neutrinos), e.g. [222], and neutrino emission from our own galaxy has been suggested as explanation, e.g. [223]. A similar trend has been reported by a third-party analysis [220] of the 4-year dataset of High Energy Starting Events [182]. The authors of [220] emphasize the negligible significance due to large uncertainties, especially in the northern sky where Earth's absorption at high energies suppresses the already small neutrino flux even further. Since in this work, we have significantly lowered the energy threshold and thus observe a larger number of astrophysical neutrino events from the northern sky, we are in a unique position to study the possibility of such aa asymmetry. Finally, thanks to our extended and much improved treatment of detector-related systematic uncertainties, we will be able to show that at least for cascade-type events, different assumptions about photon scattering in the ice (bulk-ice and/or hole-ice) can make an isotropic flux appear asymmetric with respect to the two hemispheres in the detector.

The best-fit values for this 2-hemisphere model are shown in Tab. 6.10. The spectral

Parameter		Calibration	Result $\pm 1\sigma$
norm astro (north)	$\boldsymbol{\phi_N}$	-	$(\mathbf{1.15^{+0.77}_{-0.82}})$ c.u.
spectral index (north)	$\boldsymbol{\gamma_N}$	-	$\mathbf{2.43^{+0.19}_{-0.33}}$
norm astro (south)	$\boldsymbol{\phi_S}$	-	$(\mathbf{1.55^{+0.31}_{-0.28}})$ c.u.
spectral index (south)	$\boldsymbol{\gamma_S}$	-	$\mathbf{2.53 \pm 0.12}$
norm conv	ϕ_{conv}	-	$(1.01^{+0.13}_{-0.11}) \times \Phi_{HKKMS06}$
norm prompt	ϕ_{prompt}	-	$2.51^{+5.18}_{-2.51} \times \Phi_{BERSS}$
norm muon	ϕ_{muon}	-	1.43 ± 0.04
cosmic ray index dev.	$\Delta\gamma_{CR}$	0.00 ± 0.05	0.04 ± 0.03
scattering scale (bulk)	ϵ_{scat}	1.00 ± 0.07(*)	1.02 ± 0.03
absorption scale (bulk)	ϵ_{abs}	1.00 ± 0.07(*)	1.00 ± 0.04
scattering scale (HI)	ϵ_{abs}	1.00 ± 0.67	$1.62^{+0.21}_{-0.20}$
DOM efficiency	ϵ_{eff}	0.99 ± 0.10	$1.04^{+0.08}_{-0.08}$
$-2\log\Lambda_{sat}$	64.62		

TABLE 6.10: Best-Fit parameters for the phenomenological two hemisphere model (different single power laws in northern and southern skies). (*) nuisance parameters are anti-correlated with $\rho = -0.1$

properties measured for the two hemispheres appear well-consistent. We obtain a flux at 100 TeV of $\phi_N = (1.15^{+0.77}_{-0.82}) \cdot 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$ with spectral index $\gamma_N = 2.43^{+0.19}_{-0.33}$ for the northern sky and normalization $\phi_S = (1.55^{+0.31}_{-0.28}) \cdot 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$ with spectral index $\gamma_S = 2.53 \pm 0.12$ for the southern sky. Both appear well-consistent. This scenario can be compared to the isotropic, single powerlaw best fit (Tab. 6.1) by means of a likelihood ratio test.

$$H_0 : (\phi_N,\, \gamma_N) = (\phi_S,\, \gamma_S) \quad \text{against} \quad H_1 : (\phi_N,\, \gamma_N) \neq (\phi_S,\, \gamma_S) \tag{6.20}$$

This test satisfies all conditions of Wilks' theorem, eq. (5.26) (no boundary effects, no parameters undefined under H_0), and thus we find $p = 0.82$ for the observed test-statistic value of $-2\log\Lambda = 0.41$. We therefore conclude that there is no evidence for tension between the astrophysical spectra inferred separately for the two hemispheres. Fig. 6.35 shows the two-dimensional 68% confidence contours for the astrophysical parameters inferred from the two hemispheres, northern sky (solid black line and black circle) and southern sky (solid red line and red circe), in comparison to the baseline result assuming an isotropic flux (blue solid line, blue star). The same figure also shows other solutions (dashed and dotted lines) that we will come back to later in this section. The flux is measured much more precisely in the southern sky than in the northern sky. The difference in precision is too large to be explained by the differences in the total number of observed events at high energies. As discussed in several places, e.g. Sec. 6.3 and Sec. 6.5, due to the absence of any atmospheric self-veto capabilities in the north, there exists a large degeneracy between the spectral shape of the prompt flux ($\gamma \sim 2.7$) and the rather soft astrophysical neutrino flux $\gamma_{astro} \sim 2.5$ obtained in this work. This degeneracy is reflected in the uncertainties of the flux measurement in the northern sky. In fact, the preferred solution in this two-hemisphere model finds a non-zero prompt component with normalization $\Phi_{prompt} = 2.51^{+5.18}_{-2.51} \times \Phi_{BERSS}$. Given the large uncertainty, which makes this solution well consistent with zero prompt flux, we can not claim, that in this scenario we would have observed a prompt atmospheric neutrino flux. This large degeneracy between

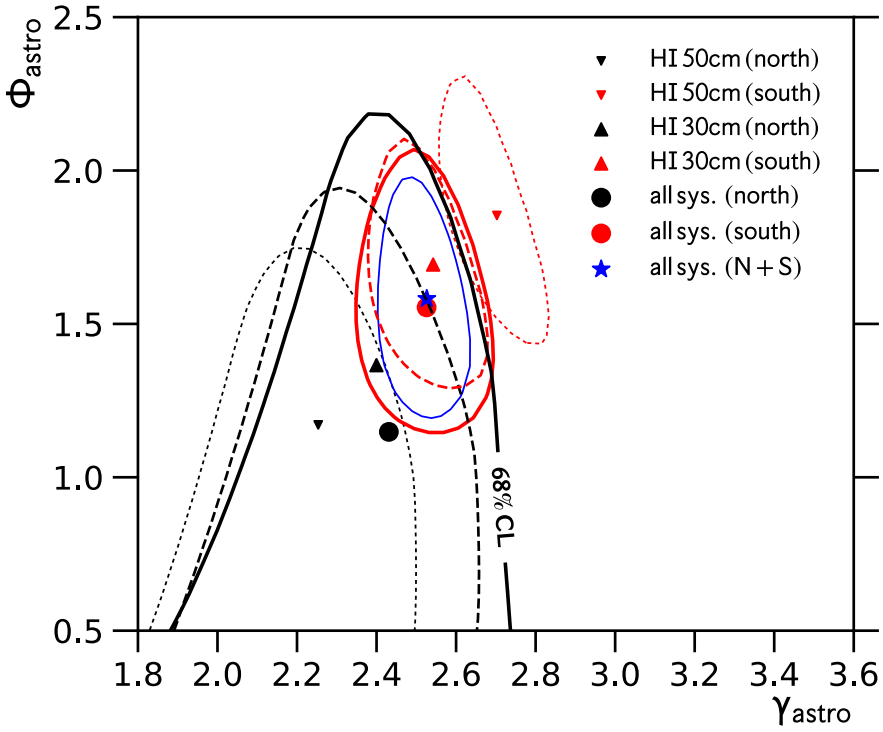


FIGURE 6.35: Confidence contours for parameters in two hemisphere model. Northern sky (black) and southern sky (red). Comparison to isotropic flux assumption (blue). Various assumptions about detector systematics (see text).

both components strongly impedes the ability of this measurement to disfavor by itself large spectral differences in the astrophysical flux from both hemispheres. In particular, this measurement, while consistent with an isotropic flux, would also be consistent with an astrophysical flux in the northern hemisphere with spectral properties as different, i.e. as hard, as the one observed with muon neutrinos from northern sky at higher energies [205]. However, a previous iteration of the same analysis [201] obtained a strong upper limit for the prompt atmospheric neutrino flux $\Phi_{prompt} < 1.06 \cdot \Phi_{ERS}$, which is at the level of the best-fit prompt normalization for this 2-hemisphere scenario (see Tab. 6.10 and Sec. 6.5 for details). Fig. 6.36 (top left) shows how the conditional best-fit value for the prompt normalization depends on the assumed astrophysical neutrino flux in the northern sky. The problematic degeneracy is shown as an increase in prompt normalization in the direction of smaller astrophysical normalization with harder spectral indices (i.e. in the direction of the best-fit values for the muon neutrino measurement [205]). To better facilitate a comparison between both measurements in the northern sky, we repeated the 2-hemisphere fit under the assumption that the prompt normalization is vanishingly small ($\Phi_{prompt}/\Phi_{BERSS} = 0$), consistent with the best-fit and strong upper-limit reported in [201]. This breaks the degeneracy. The result is shown in Fig. 6.36 (bottom) as the dashed (dotted) black line, while the result of [205] is shown in pink for comparison. The best-fit values are marked by up-triangles and given in Tab.6.11. Compared to the best-fit (Tab. 6.10, black and red points in Fig. 6.36 (bottom)), forcing the prompt normalization to 0 improves the consistency between both hemispheres event further, while the quality

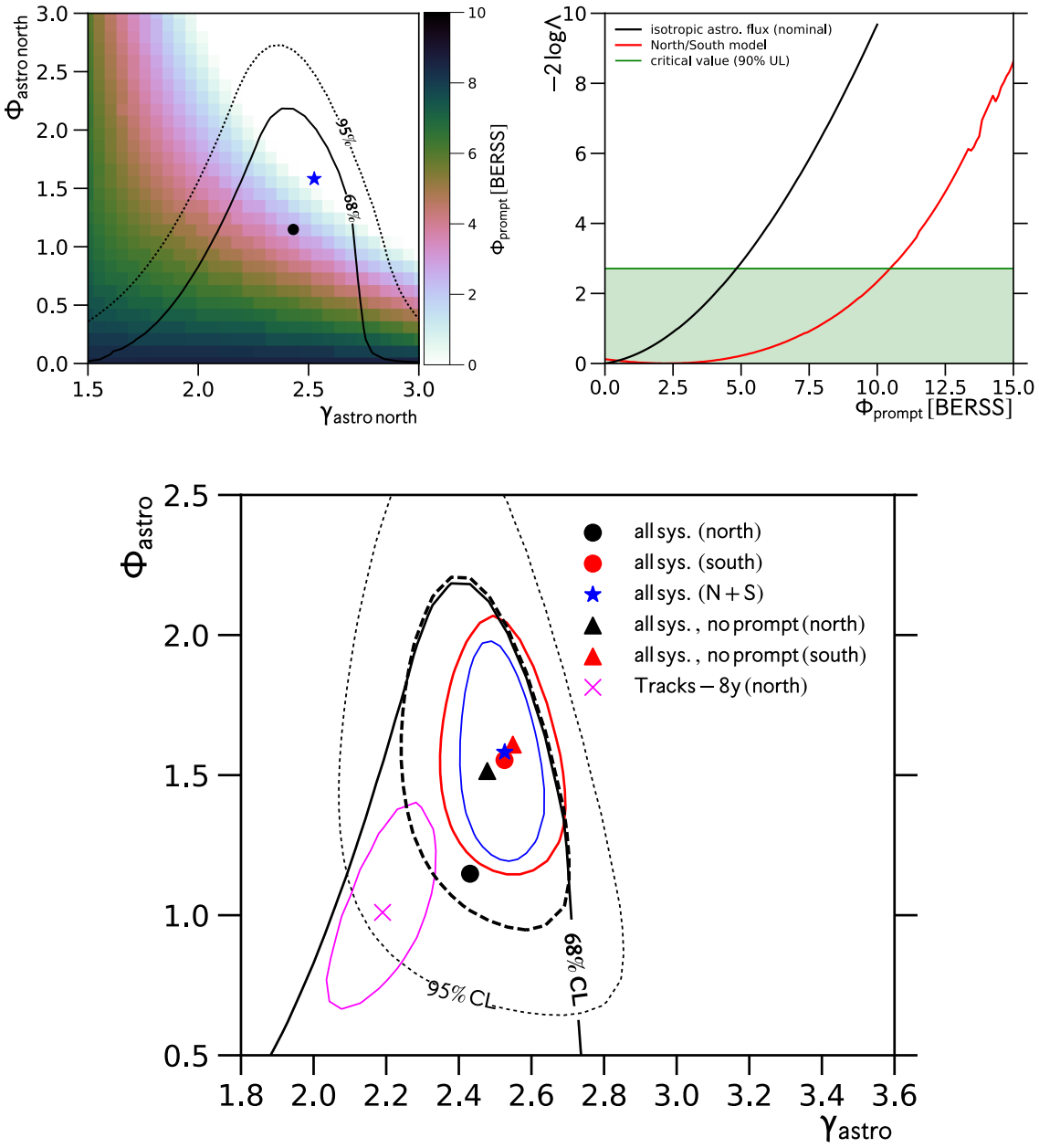


FIGURE 6.36: Top left: northern sky flux measurement assuming no prompt flux exists. Top right: constraints on prompt flux normalization if astrophysical flux is allowed to be different in the two hemispheres. Bottom: Confidence contours for parameters in two hemisphere model with (solid) and without prompt flux (dashed, dotted)). Northern sky (black) and southern sky (red). Comparison to isotropic flux assumption (blue).

of the fit ($-2\log\Lambda = 0.13$) is essentially identical. Since the degeneracy is broken, the uncertainties for the astrophysical flux parameters (Northern Sky) are much reduced. They remain, however, larger than the ones obtained in the southern sky because of the larger conventional atmospheric neutrino background at low energies and smaller number

Parameter		Calibration	Result $\pm 1\sigma$
norm astro (north)	$\phi_{\mathbf{N}}$	-	$(\mathbf{1.51}^{+\mathbf{0.21}}_{-\mathbf{0.39}})$ c.u.
spectral index (north)	$\gamma_{\mathbf{N}}$	-	$\mathbf{2.48}^{+\mathbf{0.15}}_{-\mathbf{0.15}}$
norm astro (south)	$\phi_{\mathbf{S}}$	-	$(\mathbf{1.61}^{+\mathbf{0.30}}_{-\mathbf{0.26}})$ c.u.
spectral index (south)	$\gamma_{\mathbf{S}}$	-	$\mathbf{2.55 \pm 0.10}$
norm prompt	ϕ_{prompt}	-	0.0 (fixed)
$-2\log\Lambda_{sat}$		64.75	

TABLE 6.11: Best-Fit values for two-hemisphere model assuming no prompt flux exists.

Parameter		Calibration	Result $\pm 1\sigma$
norm astro (north)	$\phi_{\mathbf{N}}$	-	$(\mathbf{1.37}^{+\mathbf{0.38}}_{-\mathbf{0.86}})$ c.u.
spectral index (north)	$\gamma_{\mathbf{N}}$	-	$\mathbf{2.40}^{+\mathbf{0.16}}_{-\mathbf{0.27}}$
norm astro (south)	$\phi_{\mathbf{S}}$	-	$(\mathbf{1.69}^{+\mathbf{0.27}}_{-\mathbf{0.26}})$ c.u.
spectral index (south)	$\gamma_{\mathbf{S}}$	-	$\mathbf{2.54}^{+\mathbf{0.09}}_{-\mathbf{0.11}}$
scattering scale (HI)	ϵ_{abs}	-	1.67 (fixed)
$-2\log\Lambda_{sat}$		67.89	

TABLE 6.12: Best-Fit values for two-hemisphere model assuming a 30 cm scattering length in IceCube's hole ice. All other detector-related systematics parameters are fixed to their nominal values (see text).

of events in the background free regime at higher energies). While still consistent with the muon neutrino measurement at the $< 2\sigma$ level, our Northern-Sky flux measurement now rejects the corresponding best-fit astrophysical flux parameter values at the $\sim 90\%$ confidence level. This demonstrates that introducing a North-South asymmetry into the astrophysical spectrum does not appear to meaningfully improve the agreement between both measurements (other than increasing uncertainties).

Impact of Systematic Uncertainties

How would the conclusions differ, had we not introduced an improved treatment of systematic uncertainties into the analysis? To answer this question, we redo the fit under different assumptions. First, we ignore detector-related systematic uncertainties but assume a detector with a hole-ice scattering length of 30 cm that otherwise behaves nominally. The result is shown in Fig. 6.35 for both hemispheres: Northern Sky (black dashed contour) and southern sky (red dashed contour). The best-fit values are marked by up-triangles and are given in Tab. 6.12. The quality of the fit worsens by 3.3 units of $-2\log\Lambda$, but remains good. The change is not significant, since we forced four parameters to their nominal values. The agreement between both hemispheres, while still good, also becomes slightly worse. This can be evaluated by re-doing the relevant test, eq. (6.20). The relevant isotropic single powerlaw fit result corresponding to the identical assumptions about systematic uncertainties is the green contour of Fig. 6.9 (bottom left). The test-statistic value for a test between that result and the one in Tab. 6.12 is $-2\log\Lambda = 2.58$, which is larger than what we obtained for the full treatment of systematic uncertainties ($-2\log\Lambda = 0.41$). Correspondingly, we find a smaller p-value of $p = 0.3$, which still does not constitute any interesting evidence for a North-South asymmetry in the flux.

The second assumption is more extreme. We keep all detector systematics related parameters at their nominal values, as before. However, in addition, we also assume a nominal

Parameter		Calibration	Result $\pm 1\sigma$
norm astro (north)	$\phi_{\mathbf{N}}$	-	$(\mathbf{1.17}^{+\mathbf{0.40}}_{-\mathbf{0.76}})$ c.u.
spectral index (north)	$\gamma_{\mathbf{N}}$	-	$\mathbf{2.25}^{+\mathbf{0.16}}_{-\mathbf{0.25}}$
norm astro (south)	$\phi_{\mathbf{S}}$	-	$(\mathbf{1.85}^{+\mathbf{0.31}}_{-\mathbf{0.27}})$ c.u.
spectral index (south)	$\gamma_{\mathbf{S}}$	-	$\mathbf{2.70}^{+\mathbf{0.09}}_{-\mathbf{0.10}}$
scattering scale (HI)	ϵ_{abs}	-	1.00 (fixed)
$-2\log\Lambda_{sat}$		83.19	

TABLE 6.13: Best-Fit values for two-hemisphere model assuming a 50 cm scattering length in IceCube's hole ice. All other detector-related systematics parameters are fixed to their nominal values (see text).

hole-ice scattering length of $\lambda = 50\,\text{cm}$, which has been the baseline assumption in previous IceCube analyses. Repeating the fit, we obtain the dotted contours shown in Fig. 6.35, which show a clear separation of the measured fluxes in the two hemispheres. The best-fit values are marked as down triangles and are given in Tab. 6.13. In particular we find a quite hard spectrum for the Northern Sky with $\gamma_N = 2.25$. The fit-quality further deteriorates by 18.58 units of $-2\log\Lambda$, which corresponds to a goodness-of-fit p-value of 7.5% for this scenario. While small, based on this value alone, one would not reject this fit result. Would this fit have yielded evidence for an asymmetric flux? We apply the hypothesis test of eq. (6.20). The relevant isotropic single-powerlaw fit result for the same assumptions about the detector is now given by the red contour in Fig. 6.9 (bottom left). The observed test-statistic value would become $-2\log\Lambda = 21.62$ corresponding to a p-value of $2 \cdot 10^{-5}$, or equivalently $4.1\,\sigma$[12]. This would strongly favor the presence of a north-south asymmetry in the astrophysical neutrino flux, while not necessarily being flagged as a bad fit by the more general fit quality criteria.

Impact on Prompt Upper-Limit

As discussed in Sec. 6.5 the ability of this analysis to constrain the intensity of the prompt atmospheric neutrino flux partially depends on the assumption that the astrophysical neutrino flux is isotropic, in particular symmetric with respect to both hemispheres. The 2-hemisphere astrophysical flux model can be used to check, by how much the limit would weaken if we were to relax this condition. Fig. 6.36 (top right) shows the profile-likelihood function for the prompt atmospheric neutrino flux normalization using the 2-hemisphere model for the astrophysical neutrino flux (red) in comparison to the nominal isotropic assumption (black). We observe significant weakening of the 90% C.L. upper limit from $4.8 \cdot \Phi_{BERSS}$ to $10.5 \cdot \Phi_{BERSS}$ or, equivalently, $3.7 \cdot \Phi_{ERS}$. Thus far, no IceCube analysis has obtained evidence for potential anisotropies in the flux of high energy astrophysical neutrinos, neither in the track [87][62] nor cascade [224] channel. Thus, assuming an extra-galactic origin with isotropic arrival directions appears reasonable.

[12]This is assuming that Wilk's theorem ($k = 2$) holds far into the tail of the test-statistic distribution. One would check this by means of simulated toy-experiments.

Chapter 7

Summary and Outlook

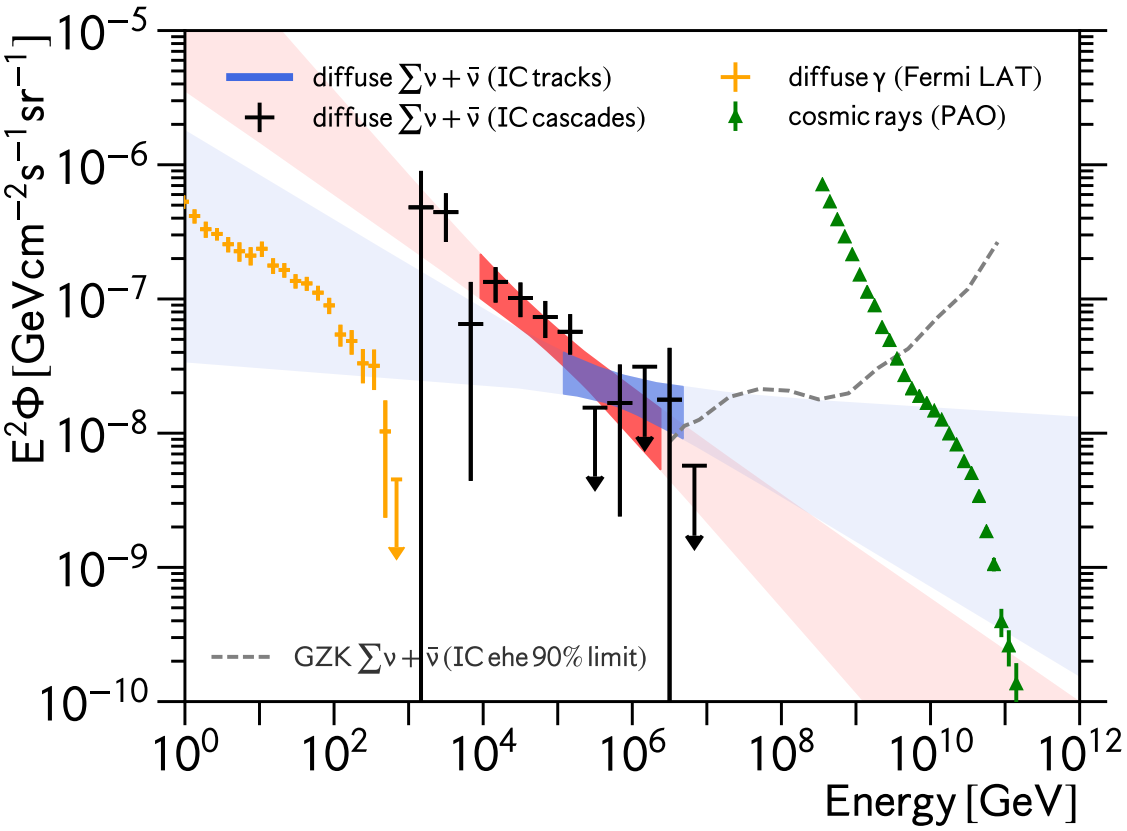


FIGURE 7.1: Summary of fluxes of high energy cosmic messengers: γ-rays observed by Fermi [225] (orange), cosmic-rays recorded by the Pierre Auger Observatory [226] (green), the neutrino flux measured here (red and black), the neutrino flux observed in the track channel (ν_μ) [205] an limit on GZK neutrinos (all flavors) set by IceCube [227]

In this dissertation we studied in-depth the flux of high energy astrophysical neutrinos, discovered by the IceCube experiment [5]. For this purpose we developed a new event selection focusing on neutrino induced cascades, mainly sensitive to the electron and tau

neutrino component of this flux. Compared to previous works [164][155] we increased the electron neutrino effective area at all energies, ranging from a 25% increase at PeV energies to more than a factor of 10 increase around 1 TeV. We applied the methods experimental data recorded with the IceCube detector in the years 2012 to 2015 and obtained IceCube's largest cascade sample to date. In total we find ~ 6300 neutrino candidate events with an estimated purity of $> 90\%$. The majority of these events $\sim 75\%$ are identified as cascade-like including an estimated number of ~ 400 astrophysical electron and tau neutrinos. We find these events to be well described by an isotropic, single powerlaw astrophysical neutrino flux with spectral index $\gamma = 2.53^{+0.07}_{-0.09}$ and per-flavor normalization $\Phi = 1.58^{+0.25}_{-0.28}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$ at a neutrino energy of 100 TeV. Compared to previous works we have improved the treatment of systematic uncertainties, which now account for detector related uncertainties as well as uncertainties in the atmospheric neutrino flux, such as the different hadronic interaction models. The contribution from astrophysical neutrinos is favored over a purely atmospheric origin with a statistical significance of 8.7σ. We further studied simple extensions to the single powerlaw flux model that could arise from different scenarios of astrophysical neutrino production. We have not observed evidence for such additional complexity in the flux model, but uncertainties are large. Therefore only extreme cases can be excluded, such as strong spectral hardening $\Delta\gamma > 0.6$ in a two-component powerlaw scenario at 90% confidence level, assuming that both components contribute noticeably to the observed neutrino flux. Finally, we searched for a contribution from the expected flux of prompt atmospheric neutrinos from the decays of heavy mesons (charm) in Earth's atmosphere. The non-observation of this flux corresponds to a flux upper-limit of 4.8 times the benchmark pQCD prediction we considered (BERSS [102]) at 90% confidence level.

The results obtained in this dissertation are put into context with the neutrino flux measurement by IceCube in the track channel (ν_μ), and measurements of other cosmic messenger particles (γ-rays by Fermi-LAT, cosmic-rays by PAO) in Fig. 7.1. Future IceCube measurements will investigate whether the small discrepancy between the results reported here (red) and the ones measured with muon neutrinos from the Northern Sky (blue) are due to a statistical effect or due to a possible spectral hardening of the astrophysical neutrino flux above ~ 100 TeV, which could for example be indicative of an extra component in the neutrino flux. A combined analysis of the sample obtained here with 2012-2015 data and the previous cascade analysis of data recorded in 2010-2011[155]) will be performed for the publication. Additional insights will come from a global analysis of all IceCube detection channels, currently under development [202].

Considering the bigger picture, it remains to be seen whether (and if, how) the γ-ray flux measured by e.g. the Fermi satellite, the flux of cosmic-rays observed by e.g. the Pierre Auger Observatory and the flux of neutrinos detected by IceCube can be related. The intensity of the neutrino flux measured in this work (as well as in previous works) might require neutrino production in dense astrophysical environments that high energy γ-rays can not efficiently escape from [79][77]. The quest to understand the high energy, non-thermal universe continues.

Appendix A

Additional Analysis Topics

In this Chapter we will discuss the stability of the fit results against time (Sec. A.1 and provide some further insights into the preference for enhanced photon scattering in IceCube's drill holes, reported in this work (Sec. A.2).

A.1 Analysis of Individual Years of Data Taking

In the previous section we studied the spectral and, to a lesser degree, the directional properties of the astrophysical neutrino flux. Here we will briefly study whether or not the results show evidence for time dependence. This is mainly intended as a cross-check. IceCube operates and takes data on a yearly basis, the so-called "IceCube seasons", usually starting in May. Atmospheric neutrino background, while variable throughout one season ($\sim 5\%$), is not expected, and has not been observed to, vary between seasons [228]. Except for possible transient astrophysical phenomena, the astrophysical neutrino flux is expected to be largely time-independent.

First, we investigated whether the astrophysical single powerlaw parameters (Φ, γ), measured separately in each of the four seasons (2012/13/14/15), appear consistent with each other. This can be tested by modeling the atmospheric background as well as the detector systematics to be independent of the season (as before) but introducing extra parameters for the astrophysical flux in each season: (Φ_i, γ_i), with $i \in \{12, 13, 14, 15\}$. This increases the total number of astrophysical flux parameters from 2 to 8. Because of the large number of parameters, we kept detector related parameters fixed at their nominal values, except for a hole-ice scattering length of 30 cm. Fig. A.1 (top) shows the result of this fit (each color represents one season) in comparison the to the result obtained for the time-independent assumption (blue). Obviously, the uncertainties for the "yearly" astrophysical parameters are quite large, since only a fraction of the available dataset can be used to obtain parameter constraints. The best-fit value are given in Tab. A.1. The large observed value of $-2 \log \Lambda_{sat} = 231.23$ is expected. To perform the fit, we increased the total number of bins by a factor of 4 (total number of seasons) from 78 to 312. Fig. A.1 (top) does not show obvious outliers and all years appear consistent. To make this

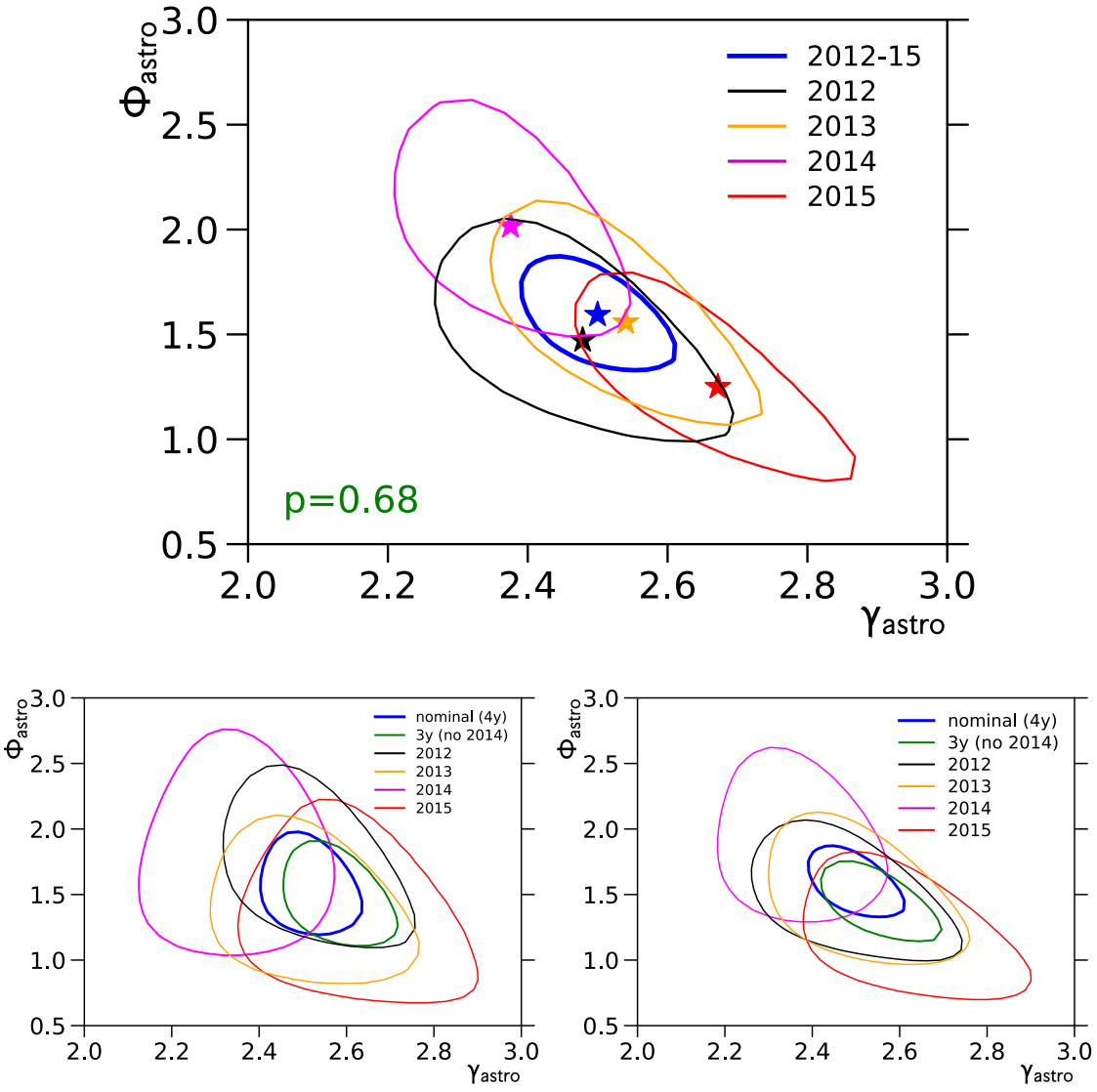


FIGURE A.1: Separating the flux measurement by year: fitting separate astrophysical fluxes in each year, but sharing background flux normalizations (top). completely separate fits with (bottom left) and without (bottom right) detector systematics.

observation quantitative, we perform the following hypothesis test:

$$\begin{aligned} H_0 &: (\Phi_i, \gamma_i) = (\Phi_j, \gamma_j) \wedge \boldsymbol{\theta_i} = \boldsymbol{\theta_j}, \forall i, j \in \{12, 13, 14, 15\} \quad \text{against} \\ H_1 &: \exists\, i \neq j,\ s.t.\ (\Phi_i, \gamma_i) \neq (\Phi_j, \gamma_j) \wedge \boldsymbol{\theta_i} = \boldsymbol{\theta_j}, \forall i, j \in \{12, 13, 14, 15\} \end{aligned} \tag{A.1}$$

where $\boldsymbol{\theta}$ denotes the vector of nuisance parameters (here: normalizations of atmospheric fluxes). Using the same binning, we find ($\Phi = 1.59$, $\gamma = 2.5$) (blue contour in Fig. A.1 (top)), which is essentially indistinguishable from the equivalent result obtained in 6.9 (bottom left, green contour). The corresponding observed value for the likelihood ratio test-statistic is $-2 \log \Lambda = 3.96$. Since the conditions for Wilks' theorem are satisfied, we find $p = 0.68$. Thus, the observed behavior across the 4 years of data taking is

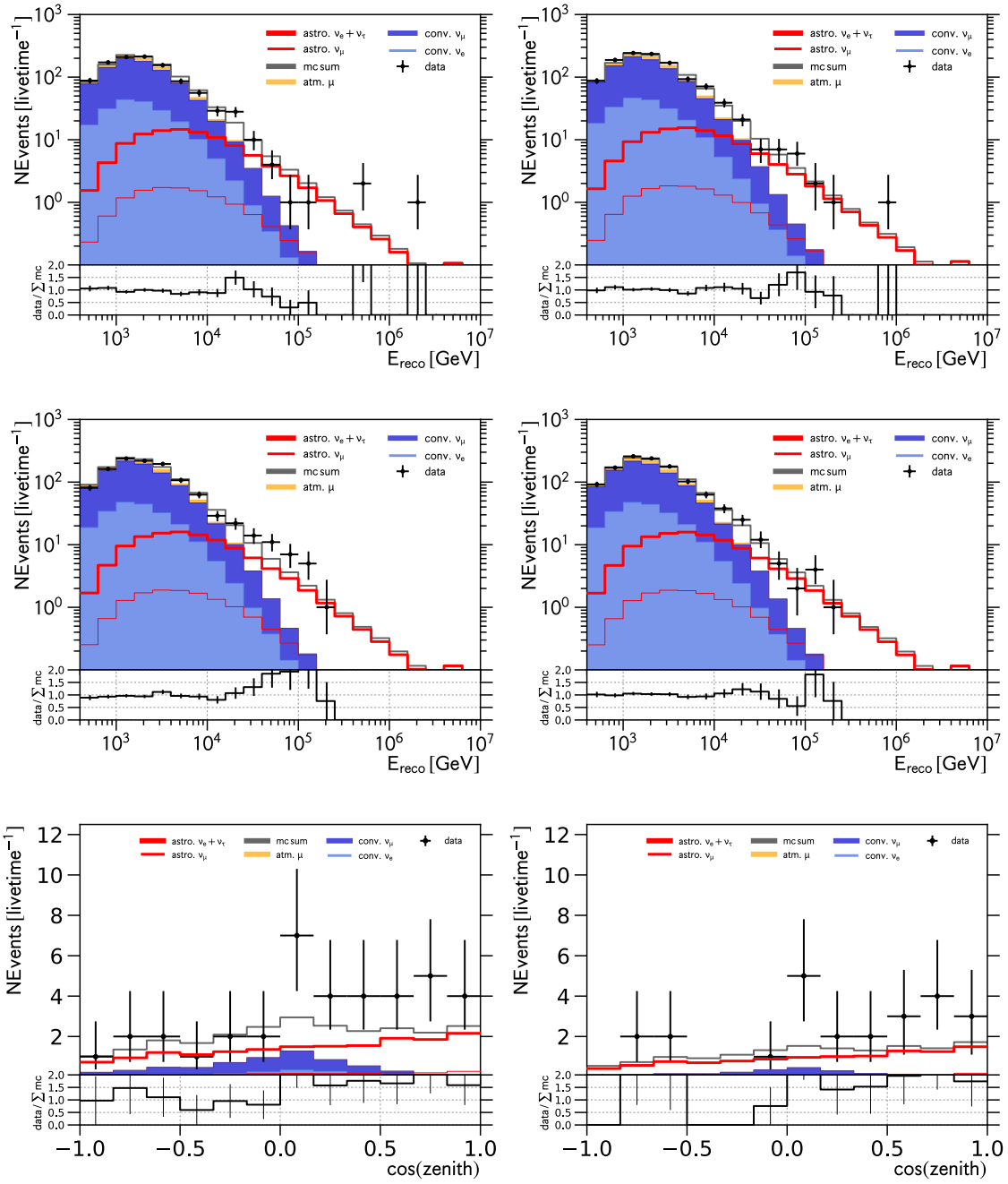


FIGURE A.2: Cascade energy spectral (all zenith angles) observed in the four years: 2012 (top left), 2013 (top right), 2014 (middle left) and 2015 (middle right). Bottom: zenith distribution for year 2014 with $E > 25\,\text{TeV}$ (left) and $E > 40\,\text{TeV}$ (right)

Parameter		Calibration	Result $\pm 1\sigma$
norm astro (12)	ϕ_{12}	-	$(\mathbf{1.47^{+0.37}_{-0.33}})$ c.u.
spectral index (12)	γ_{12}	-	$\mathbf{2.48^{+0.14}_{-0.14}}$
norm astro (13)	ϕ_{13}	-	$(\mathbf{1.56^{+0.38}_{-0.34}})$ c.u.
spectral index (13)	γ_{13}	-	$\mathbf{2.54 \pm 0.13}$
norm astro (14)	ϕ_{14}	-	$(\mathbf{2.02^{+0.39}_{-0.36}})$ c.u.
spectral index (14)	γ_{14}	-	$\mathbf{2.38 \pm 0.11}$
norm astro (15)	ϕ_{15}	-	$(\mathbf{1.25^{+0.35}_{-0.31}})$ c.u.
spectral index (15)	γ_{15}	-	$\mathbf{2.67 \pm 0.13}$
norm conv	ϕ_{conv}	-	$(1.05 \pm 0.02) \times \Phi_{HKKMS06}$
norm prompt	ϕ_{prompt}	-	$< 0.9 \times \Phi_{BERSS}$
norm muon	ϕ_{muon}	-	1.47 ± 0.02
scattering scale (HI)	ϵ_{scat}	CONST.	1.67
$-2 \log \Lambda_{sat}$	231.23		

TABLE A.1: Best-fit values for the astrophysical flux parameters separated by year. Background flux parameters are shared by all years.

perfectly consistent with the same single powerlaw astrophysical neutrino flux. So far, this "year-by-year" analysis assumed that background and detector related parameters are constant (shared) between the years. This can further be relaxed by simply performing the entire analysis (c.f. Sec. 5) individually for each year. The result is shown in Fig. A.1 (bottom left: no detector systematic uncertainties; bottom right: with detector systematic uncertainties). The conclusions remain unchanged. The observed best-fit spectra, assuming the combined single powerlaw best-fit values (Tab. 6.1), are shown in Fig. A.2 (all cascades) for the years 2012 (top left) to 2015 (2nd row, right). The same spectra, when separated into the three declination bins, can be found in Appendix A: Fig. A.6 (up-going cascades) and Fig. A.7 (near vertical (top) and horizontally down-going (bottom) cascades). In general the simulation prediction describes the data observed in the different periods well. However we found correlated excess of data events over expectations in the year 2014 at high energies between 25 TeV and 250 TeV that extends over 5 consecutive bins and requires further study.

Southern Hemisphere High Energy Excess in 2014

To learn about the observed excess, we attempted to correlate the corresponding high energy events with other observables to identify possible common characteristics. These include among others the cut-variables, reconstructed location of the events within the detector, arrival direction, total PMT charge, BDT classification scores, etc. While the events in question appear to stem from the southern hemisphere, no other clustering has been found. The preference for the southern sky is shown in Fig. A.2 (bottom) for two different energy thresholds: $E > 25\,\text{TeV}$ (left) and $E > 40\,\text{TeV}$ (right). The excess is quite evenly distributed across $\cos\theta_{reco} > 0.0$ and does not appear to further prefer a certain declination. Additional analysis showed that the events also do not cluster in right-ascension and thus do not prefer a certain location in the sky[1]. The number of events from the southern hemisphere with energies within the questionable energy range

[1]This does not necessarily fully exclude a common origin, due to the rather poor angular resolution of cascades.

for each year are as follows: 10 (2012), 11 (2013), **28 (2014)** and 14 (2015). Such asymmetric behavior is not expected from any of the non-time dependent explanations one might consider: underestimated cosmic ray muon background, miss-modeling of atmospheric neutrino background including the atmospheric neutrino self veto effect, other detector related systematics, and probably also a cosmic origin due to the absence of any corresponding hot-spot in the sky.
In order to exclude the possibility of problems in the data processing chain for this particular region of observable space in the year 2014, we isolated events before any neutrino selection criteria (BDT, high energy cuts) were applied. The selection corresponds to all pre-selection criteria (L3A-D + L4A). In addition we required the events to be reconstructed as down-going $\cos\theta_{reco} > 0.0$ and the total charge collected by the PMTs to be larger than $Q = 1000\,\text{p.e.}$ to match the total PMT charge of the excess events. Roughly

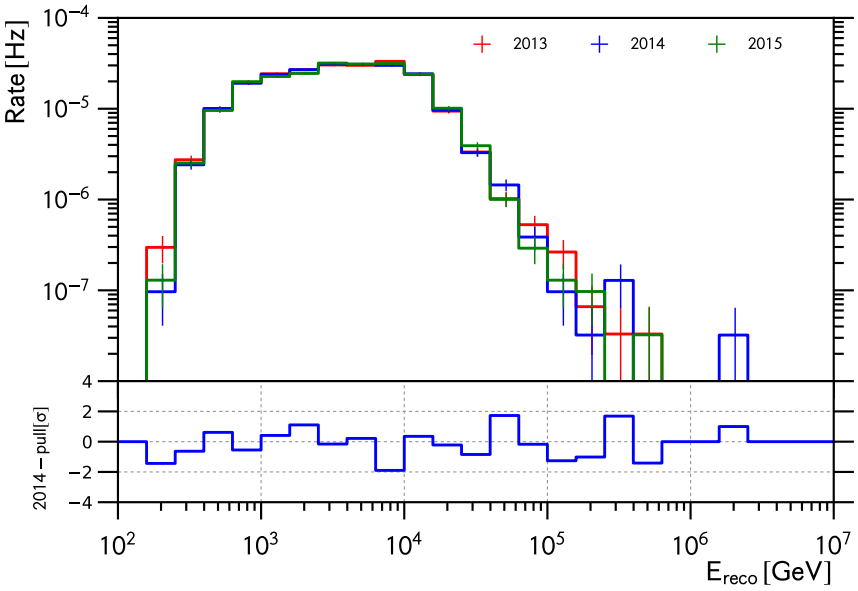


FIGURE A.3: Comparison of the data rates (2013-2015) in the background region of events with similar properties to the ones contributing to the excess in 2014.

700,000 cosmic-ray muon events pass these criteria in each year. The contribution from neutrinos is negligible ($< 0.5\%$) for this selection. The reconstructed energy distribution, obviously dominated by miss-reconstructed events, is shown in Fig. A.3 for the year 2014 (blue) and the adjacent years 2013 (red) and 2015 (green). The same figure also shows for each energy bin the observed pull (in units of significance σ) of the number of events in 2014 from the average expectation given by the other two years. No (correlated) excess is observed and the background rates after all pre-filtering criteria are stable. There are additional reasons that disfavor muon background to be responsible for the excess:

- The excess is distributed across the entire fiducial volume without preference for the detector boundaries.
- The excess persists even for the highest cascade.score classification (highest quality events).

- Manual inspection of the events did not reveal any muon-background-like characteristics of the events. Their appearance/signature is well-consistent with expectations from neutrino interactions.[2]

What is the significance of the excess? As we will show, the answer is highly dependent on the calculation, in particular how much it is tuned to the actual observation. One model-dependent possibility is to first fit the "standard" model (single powerlaw astrophysical neutrino flux, atmospheric background fluxes and detector systematics as before, described in Sec. 5) to all data excluding the year 2014, in order to obtain a baseline expectation. This results in the green contour in Fig. A.1 (bottom right). Subsequently one would calculate the probability of observing at least 28 events, taking into account the uncertainties determined in the previous step. We calculated the corresponding posterior predictive distribution [221] for observing n_{14} events in the excess region r ($25\,\mathrm{TeV} < E < 250\,\mathrm{TeV}$ and $\cos\theta > 0.0$) with expectation $\mu_r\,(\gamma, \Phi)$:

$$p\left(n_{14}\,|\,\mathbf{X}_{12/13/15}\right) = \int d\gamma \int d\Phi\, p\left(n_{14}\,|\,\mu_r\,(\gamma,\,\Phi)\right) f\left(\gamma,\,\Phi\,|\,\mathbf{X}_{12/13/15}\right) \tag{A.2}$$

where $f\left(\gamma,\,\Phi\,|\,\mathbf{X}_{12/13/15}\right)$ denotes the joint posterior distribution for γ and Φ obtained from a fit to the 2012/13/15 data $\mathbf{X}_{12/13/15}$. The result is shown in Fig. A.4 (top left). We find $p\left(n_{14} \geq 28\,|\,\mathbf{X}_{12/13/15}\right) = 4 \cdot 10^{-4}$, i.e. a significance of 3.4σ. However, since we have no alternative (and testable) model in mind that would better predict this observation, we are inclined to interpret this as an unlikely fluctuation, rather than as evidence for an unidentified time-dependent process. For one, we obtain lower significances, if we search for discrepancies in general, rather than to look for a specific correlated excess (which we only did, because we observed one.). For example, using the saturated poisson likelihood goodness-of-fit measure from eq. (5.37), we find that the following posterior predictive goodness-of-fit probabilities (Fig. A.4, 2nd row): $p = 0.014$ (left) for the 5 excess bins (southern sky w. $\cos\theta > 0.0$, 25 TeV to 250 TeV) and $p = 0.15$ (right) for evaluating all bins and thus measuring the compatibility between all observed cascade events in 2014 and the model prediction obtained from fitting the other 3 years.
Second, we had no a-priori reason to group the years in the way we did above, i.e. to declare the year 2014 as different from the others. The impact of this choice can be evaluated from the data without reference to any detector and flux models, by only assuming time-independence of the rates R_r in the excess region ($25\,\mathrm{TeV} < E < 250\,\mathrm{TeV}$ and $\cos\theta > 0.0$) as follows:

$$H_0 : R_r^{12} = R_r^{13} = R_r^{14} = R_r^{15} \equiv R_r \tag{A.3}$$

where the rate R_r^j relates to the observed number of events n_j via the detector livetime τ_j of each year.

$$n_j \sim \mathrm{poisson}\left(R_r^j \cdot \tau_j\right) \tag{A.4}$$

[2] One neutrino event (cascade) shows signs of a couple of early hits at the top of the detector, consistent with being accompanied by a low energy muon, which indicates a common atmospheric origin. We identified another event to be a miss-classified starting track and thus to be a muon neutrino. A third event (cascade) appears coincident with an uncorrelated atmospheric muon that creates hits in a different part of the detector at a much later time within the same trigger window. All three are consistent with expectations.

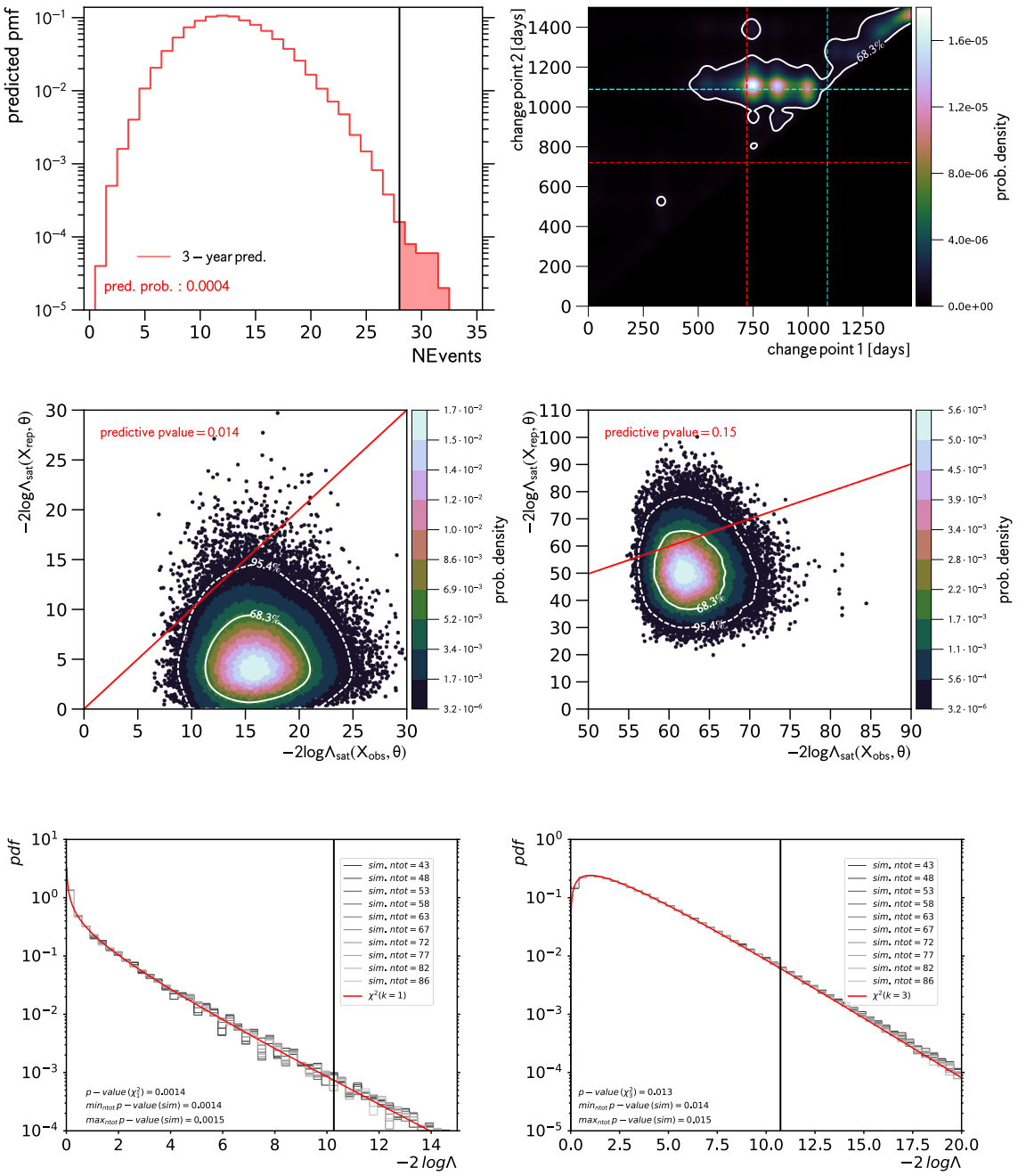


FIGURE A.4: Various possibilities to a-posteriori quantify the observed discrepancy in the data taking year 2014. see text for details.

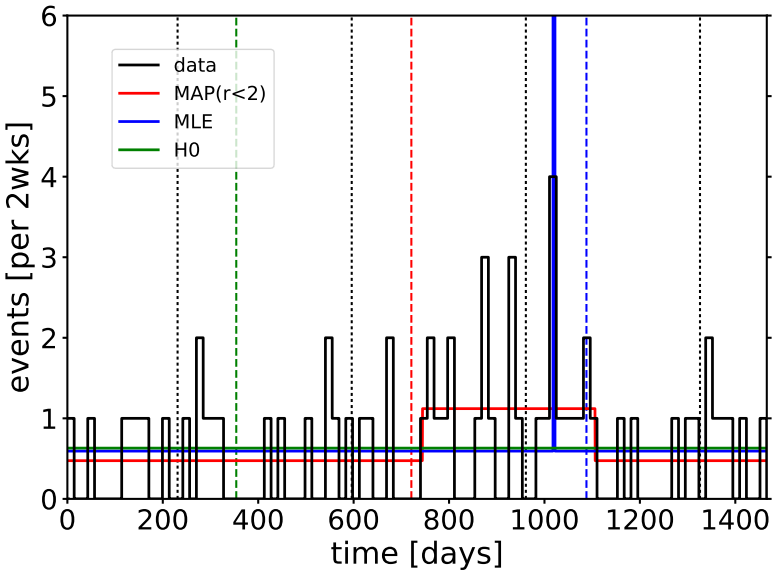


FIGURE A.5: Time series of events with high energies ($E > 25\,\text{TeV}$) from the southern sky throughout the four years (2012-2015). Transitions between calendar years are marked by the vertical dotted lines. Transitions between detector seasons are marked by vertical dashed lines (colors). Horizontal lines mark the solutions of possible change-points models (see text).

Two alternative hypothesis could be used.

$$H_1 : R_r^{14} \neq R_r^{12/13/15}, \quad R_r^{12} = R_r^{13} = R_r^{15} \equiv R_r^{12/13/15} \tag{A.5a}$$

$$H_1 : \exists\, i \neq j \;\; s.t. \;\; R_r^i \neq R_r^j,\; i, j \in \{12, 13, 14, 15\} \tag{A.5b}$$

The difference becomes apparent when the number of parameters is considered. Eq. (A.5a) requires 2 parameters (one rate for 2014 and one rate for the other years), while eq. (A.5b) specified 4 parameters (one for each year). Applying both tests to the observed data and calculating the test-statistic distributions from simulations (see Fig. A.4 bottom), we find $p = 0.0015$ using eq. (A.5a)) and $p = 0.015$ using eq. (A.5b)) - a factor of 10 difference. In the latter case, while not indicating good agreement, we would be much less confident in claiming a possible time-dependence or inconsistency across the years. The latter case, however, is the one, that appears more reasonable if we were to truly consider the possibility that the rate could vary as a function of time.

Time Series of High Energy Events (Southern Hemisphere)

The best way to check for time-dependence in the astrophysical neutrino flux (in our data sample to be more precise), is to study the time series of the events in the excess region (high energy $E > 25\,\text{TeV}$ and down-going: $\cos\theta > 0.0$). Fig. A.5 shows the number of observed cascade events with these properties per 2-week time period. Vertical dashed lines (colors) denote the transitions from one IceCube data taking season to another. The dotted black line mark the transitions between the calendar years. Also shown are the results of various model assumptions (solid lines, colored) that are discussed below.

We define two simple models for the time-dependence of the rates (R_r^j) of high energy, southern sky events throughout the total data taking period of 4 years. The first is the

baseline model and assumes a constant rate:

$$R_r^0(t) = r_0 \tag{A.6}$$

Since we are studying an excess of events, we introduce a change-point model, which consists of up to 3 time segments. The second segment is allowed to have a different rate than the other two. Hence this model could represent an isolated period of enhanced (or reduced) neutrino detections, for example corresponding to some hypothetical astrophysical phenomenon, e.g. some object in a state of enhanced neutrino emissivity. The model reads

$$R_r^1(t) = \begin{cases} r_1, & t < t_1 \\ r_2, & t_1 \leq t \leq t_2 \\ r_1, & t_2 < t \end{cases} \tag{A.7}$$

where r_1, $r_2 > 0$ are continuous rate parameters and t_1, t_2 are two discrete parameters for the location of the change-points satisfying $t_{start} \leq t_1 < t_2 \leq t_{end} + 1$. The analysis is performed using one time bin for each day within the 4 year period (1466 days). The number of events observed for any given day follows a poisson distribution. Thus the likelihood of eq. (5.15) applies.

For the constant model we find a best-fit rate of $r_0 = 0.63 \cdot (2\,wks)^{-1}$. This solution is shown in Fig. A.5 (green solid line). Assuming the change-point model we obtain the following best-fit values:

$$t_1 = 1018\,d, \quad t_2 = 1022\,d, \quad r_1 = 0.59 \cdot (2\,wks)^{-1}, \quad r_2 = 11.2 \cdot (2\,wks)^{-1} \tag{A.8}$$

This solution, especially the large rate r_2, corresponds to the observation of 4 events that we find within a narrow time-period of 5 days (in the year 2014) and is shown in Fig. A.5 as well (blue solid line)[3]. Using a simple argument, it is possible to show that this is not significant. The time-series can be split into ~ 293 distinct 5-day intervals. Assuming a constant rate $r_0 = 0.63 \cdot (2\,wks)^{-1} = 0.23 \cdot (5\,d)^{-1}$ the probability of observing less than 4 events in one period is $p(n < 4) = \sum_{k=0}^{3} p_{pois}(k \,|\, r_0) = 0.99991 \approx 1$. The probability of observing 4 or more events in (at least) one of the periods is $p_{obs} = 1 - p(n < 4)^{293} = 0.026$. Thus in $\sim 2.6\%$ of the cases we expect to observe a localized clustering of events in time similar (or more significant) to the one shown in Fig. A.5. Since the change-point model allows for other configurations that could be preferred over the baseline model (more extended clustering) this chance probability can be interpreted as an approximate lower bound to the p-value. Thus the best-fit values for the change-point model do not provide evidence against and are consistent with the baseline model.
While no evidence has been found, one might still wonder whether we could estimate a more extended time-period if we restrict parameter r_2 to smaller values. For the constraint $r_2 < 2 \cdot (2\,wks)^{-1}$ we find

$$t_1 = 744\,d, \quad t_2 = 1106\,d, \quad r_1 = 0.47 \cdot (2\,wks)^{-1}, \quad r_2 = 1.12 \cdot (2\,wks)^{-1} \tag{A.9}$$

[3]These 4 events are not localized in the Sky.

This solution is shown in Fig. A.5 as red solid line. Under this constraint, the best-fit period of enhanced event rates roughly coincides with the 2014-data taking period. The starting point, however, is not well estimated. This is shown in Fig. A.4 (top right). The solution is multi-modal and even includes the special case of not having any change-point, i.e. where the change-points t_1, t_2 are located at the boundary and thus the baseline model is recovered: $r_1 = r_0$. In summary, due to their low event rates, the high energy events ($E > 25\,\mathrm{TeV}$) observed in this sample do not allow for strong statements about possible time-dependence. More data and/or additional prior information would be needed.

A.2 Preference for Enhanced Scattering in IceCube's Drill Holes

One of the results obtained in this work concerns the IceCube detector itself. We found that nominal assumptions about the detector response to light signals do not provide a good fit to this data sample. When studying the impact of systematic uncertainties on the observables (reconstructed energy and declination) we noticed that the reconstructed

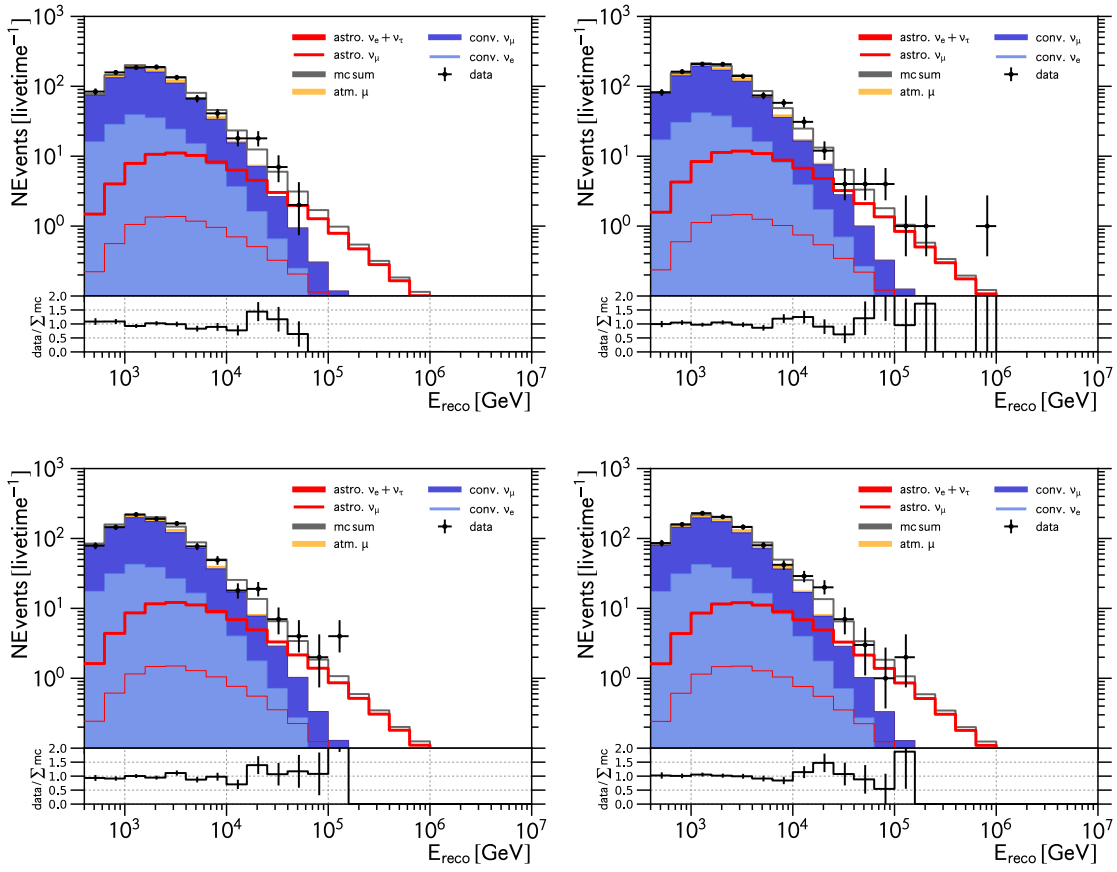


FIGURE A.6: Energy Spectra for cascades with $\cos\theta_{rec} < 0.2$ for the years 2012 (top left), 2013 (top right), 2014 (bottom left) and 2015 (bottom right)

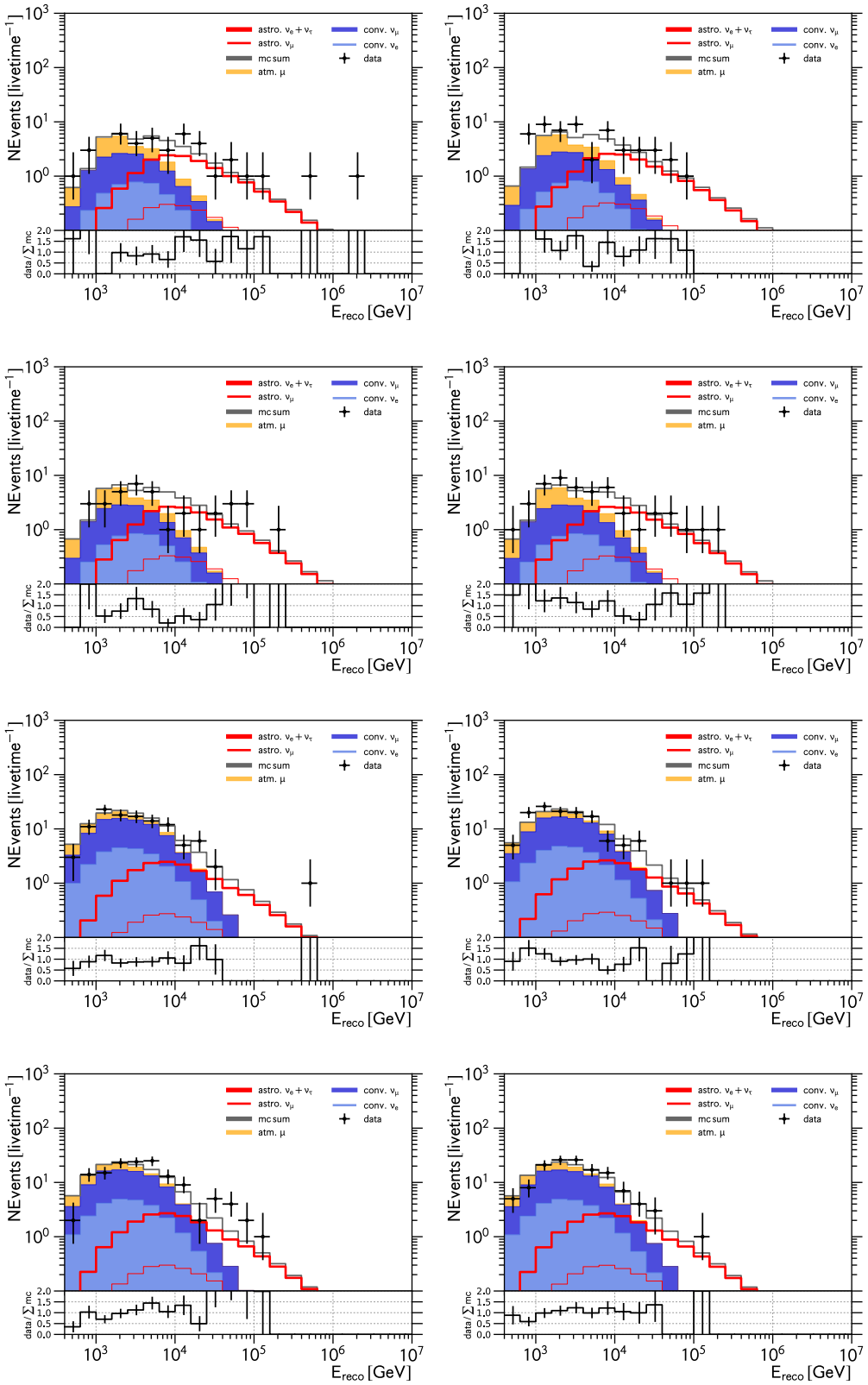


FIGURE A.7: Energy Spectra for cascades with $0.2 < \cos\theta_{rec} < 0.6$ for the years 2012 (top left), 2013 (top right), 2014 (2nd row left) and 2015 (2nd row right). Energy Spectra for cascades with $0.6 \leq \cos\theta_{rec}$ for the years 2012 (3rd row left), 2013 (3rd row right), 2014 (bottom left) and 2015 (bottom right)

$\lambda_{scat}(HI)$	$-2\log\Lambda_{sat}$ (gof)	p-value
50 cm	140.50	$9.1 \cdot 10^{-11}$
30 cm	68.87	0.032

TABLE A.2: Goodness of Fit comparison for two holeice models: hole ice scattering 50 cm and 30 cm. The fit uses 2 zenith bins, separated at $\cos\theta = 0$ (horizon).

$\lambda_{scat}(HI)$	$-2\log\Lambda_{sat}$ (gof)	p-value
50 cm	104.81	$2.8 \cdot 10^{-3}$
30 cm	70.47	0.39

TABLE A.3: Goodness of Fit comparison for two holeice models: hole ice scattering 50 cm and 30 cm. The fit uses the standard binning (3 bins with boundaries $\cos\theta \in \{-1.0, 0.2, 0.6, 1.0\}$).

zenith angle is particularly effected, see Sec. 4. Our fit result, c.f. Tab. 6.1 shows a strong preference for enhanced scattering of photons near the IceCube DOMs, i.e. a smaller scattering length of 30 cm in IceCube's drill holes. We have found in Sec. 4 that this, compared to the nominal simulation expectation based on 50, cm scattering length, enhances the expected number of events in the southern hemisphere, while it lowers the expected number of events in the northern hemisphere. The effect is predicted to be most noticeable near the horizon (c.f. Fig. 4.3, bottom) because the of the strong suppression of the expected number of events for increasingly steep trajectories (from horizontal to vertical directions). Because of this strong decrease, a small bias of a few degrees in the reconstructed zenith angle towards steeper trajectories is sufficient to effectively increase the expected number of events from the southern hemisphere. This is particularly noticeable if only 2 analysis bins, separated at $\cos\theta = 0.0$, are considered. To show the effect related to hole-ice scattering in isolation, using these 2 bins, we repeated the entire measurement of the astrophysical single powerlaw model parameters for the two cases: $\lambda = 30$ cm and $\lambda = 50$ cm, while keeping all other detector and flux related systematics parameters (dom efficiency, properties of bulk ice, cosmic-ray spectral index) fixed at their nominal values. Only flux related normalization parameters and the astrophysical spectral index are allowed to change. The result is shown in Fig. A.8 for all events (top), southern hemisphere (2nd row) and northern hemisphere (bottom). The left column assumes $\lambda = 50$ cm and the right column corresponds to $\lambda = 30$ cm. While both scenarios lead to a reasonable description of the energy spectrum for all events (top) and thus can not easily be distinguished, significant differences are observed when the two hemispheres are considered separately. The observed number of data events exceeds the best-fit expectation in the southern hemisphere by $\sim 20\%$ at all energies below ~ 150 TeV if a scattering length of $\lambda = 50$ cm is assumed (2nd row, left). Most of the excess vanishes if the value is reduced to $\lambda = 30$ cm. When the northern hemisphere is considered, instead of an excess, we observe a deficit in the number of data events compared to the expectation from the fit. The mismatch is worse for $\lambda = 50$ cm (bottom left).
Tab. A.2 quantifies this observation by evaluating the goodness-of-fit for the two cases. Simply changing the hole-ice scattering length from $\lambda = 50$ cm to $\lambda = 30$ cm improves the corresponding p-value by ~ 8 orders of magnitude.

For completeness we repeated the same calculations for the three zenith bins that we used

in the actual analysis (edges: $\cos\theta \in \{-1.0, 0.2, 0.6, 1.0\}$). The result is given in Tab. A.3. Finally, as discussed in more detail in Sec. 6, the inferred astrophysical spectral index is not effected by any of these differences (see Figs. 6.6 (top left) and 6.9 (bottom left)), while the astrophysical normalization is reduced by only 10% when scattering is increased to 30 cm,.

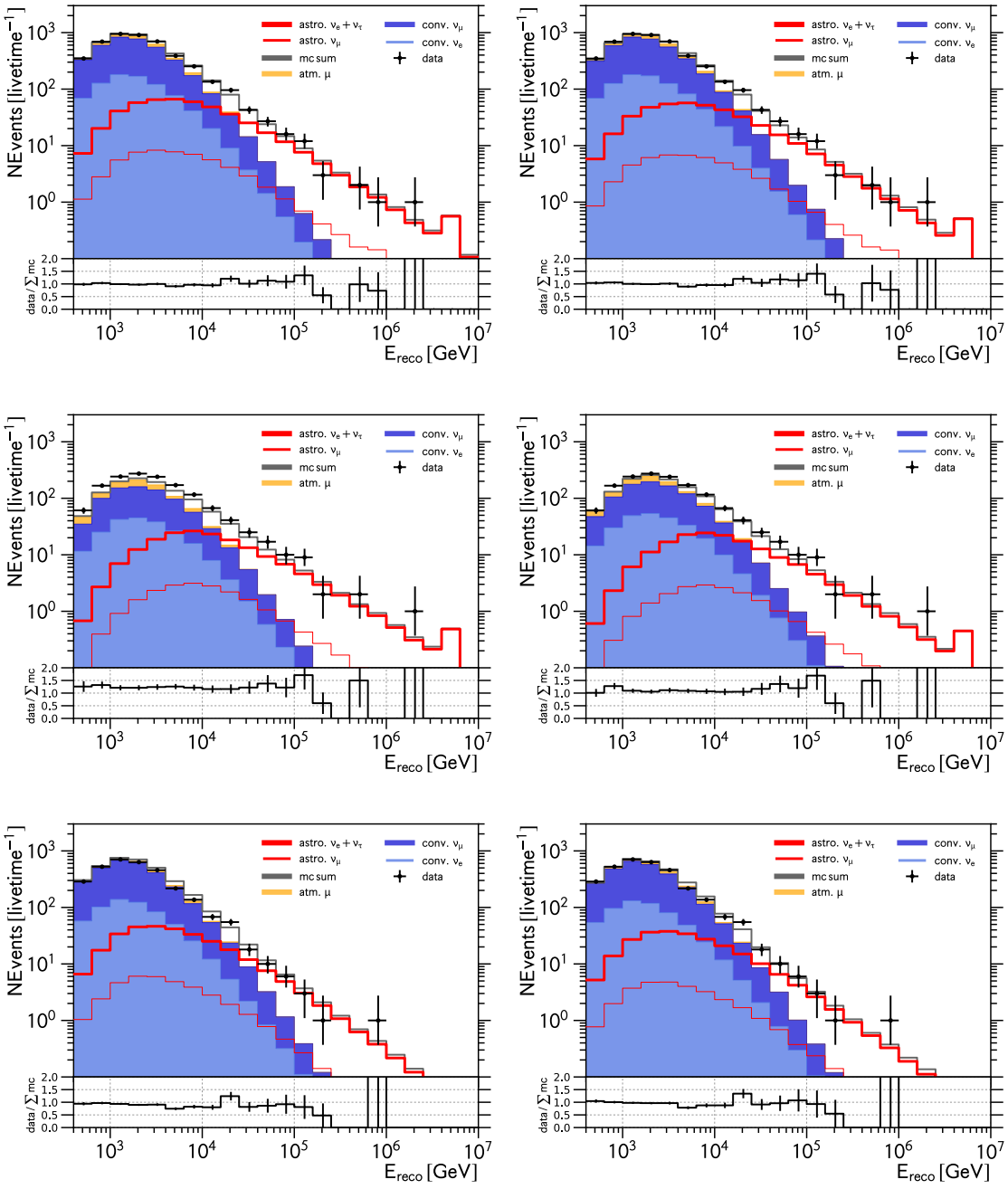


FIGURE A.8: Comparison of fitted spectra using different hole-ice models. Left: hole ice scattering 50 cm. Right: hole ice scattering 30 cm. Top: cascade from entire sky. Middle: cascades with $\cos\theta > 0$ (southern sky). Bottom: cascades with $\cos\theta \leq 0$ (northern sky).

Bibliography

[1] V. Hess. "Ueber die Bedeutung der durchdringenden Strahlung bei sieben Freiluftballonfahrten". *Physik. Zeitschrift.* XIII (1912).

[2] W. Kraushaar et al. "Explorer XI Experiment on Cosmic Gamma Rays". *Astrophysical Journal* 141 (1965).

[3] H. S. Hudson et al. "Hard Solar X-Ray Bursts Observed by OSOS-3". *Astronomical Journal* 73, 5 (1986).

[4] B. T. Cleveland et al. "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector". *The Astrophysical Journal* 496, 1 (1998).

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

[6] B. P. Abbott et al. "Observation of Gravitational Waves from a Binary Black Hole Merger". *Physical Review Letters* 116, 061102 (2016).

[7] C. Patrignani et al. "Review of Particle Physics". *Chinese Physics* C40, 100001 (2017).

[8] A. D. Panov. "Electrons and Positrons in Cosmic Rays". *Journal of Physics: Conference Series* 409 (2013).

[9] F. Fenu et al. "The cosmic ray energy spectrum measured using the Pierre Auger Observatory". *Proceedings of Science* PoS(ICRC2017), 486 (2017).

[10] D. Ivanov et al. "TA Spectrum Summary". *Proceedings of Science* ICRC 2015, 349 (2015).

[11] H. S. Ahn et al. "Energy Spectra of Cosmic-ray Nuclei at High Energies". *The Astrophysical Journal* 707, 1 (2009).

[12] D. Muller et al. "The TRACER Project: Instrument Concept, Balloon Flights, and Analysis Procedures". *Proceedings of 30th ICRC* 2, 1 (2008).

[13] O. Adriani et al. "Ten years of PAMELA in space". *LA RIVISTA DEL NUOVO CIMENTO* 10 (2017).

[14] M. Aguilar et al. "Precision Measurement of the Proton Flux in Primary Cosmic Rays from Rigidity 1 GV to 1.8 TV with the Alpha Magnetic Spectrometer on the International Space Station". *Physical Review Letters* 171103 (2015).

[15] M. G. Aartsen et al. "Measurement of the cosmic ray energy spectrum with IceTop-73". *Physical Review* D88, 042004 (2013).

[16] M. Spurio. *Particles and Astrophysics: A Multi-Messenger Approach*. 1st. Springer International Publishing, 2015.

[17] M. Ackermann et al. "Detection of the Characteristic Pion-decay Signature in Supernova Remnants". *Science* 339 (2013).

[18] P. Lagage and C. J. Cesarsky. "The maximum energy of cosmic rays accelerated by supernova shocks". *Astronomy and Astrophysics* 125, 2 (1983).

[19] T. K. Gaisser et al. *Cosmic Rays and Particle Physics*. Vol. 2nd. Cambridge University Press, 2016.

[20] E. G. Berezhko. "Maximum energy of cosmic rays accelerated by supernova shocks". *Astroparticle Physics* 5 (1996).

[21] A. M. Hillas. "The Origin of Ultra-High-Energy Cosmic Rays". *Annual Review of Astronomy and Astrophysics* 22, 1 (1984).

[22] B. Peters. "Primary Cosmic Radiation and Extensive Air Showers". *Il Nuovo Cim. XXII* (1961).

[23] W. D. Apel et al. "Kneelike Structure in the Spectrum of the Heavy Component of Cosmic Rays Observed with KASCADE-Grande". *Physical Review Letters* 107, 17 (2011).

[24] G. Giacinti et al. "Explaining the spectra of cosmic ray groups above the knee by escape from the Galaxy". *Physical Review* D90, 4 (2014).

[25] A. Aab et al. "Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8 EeV". *Science* 357, 6357 (2017).

[26] A. M. Hillas. "Cosmic Rays: Recent Progress and some Current Questions". *arXiv* astro-ph, 0607109 (2006).

[27] T. K. Gaisser. "Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio". *Astroparticle Physics* 35, 12 (2012).

[28] V. S. Berezinsky and S. I. Grigor'eva. "A bump in the ultra-high energy cosmic ray spectrum". *Astronomy and Astrophysics* 199, 1-2 (1988).

[29] J. Heinze et al. "Cosmogenic Neutrinos Challenge the Cosmic-ray Proton Dip Model". *The Astrophysical Journal* 825, 2 (2016).

[30] T. K. Gaisser et al. "Cosmic Ray Energy Spectrum from Measurements of Air Showers". *arXiv* astro-ph.HE, 1303.3565 (2013).

[31] J. G. Kirk and R. O. Dendy. "Shock acceleration of cosmic rays - a critical review". *Journal of Physics G: Nuclear and Particle Physics* 27, 7 (2001).

[32] E. Fermi. "On the Origin of the Cosmic Radiation". *Physical Review* 75, 8 (1949).

[33] M. S. Longair. *High Energy Astrophysics*. 3rd. Cambridge University Press, 2011.

[34] A. R. Bell. "The acceleration of cosmic rays in shock fronts." *Monthly Notices of the Royal Astronomical Society* 182 (1987).

[35] R. Blandford and J. P. Ostriker. "Particle acceleration by astrophysical shocks". *Astrophysical Journal* 221 (1987).

[36] L. O. Drury. "An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas". *Reports on Progress in Physics* 46, 8 (1983).

[37] L. O. Drury. "On particle acceleration in supernova remnants". *Space Science Reviews Space Science Reviews* 36, 1 (1983).

[38] L. D. Landau and E. M. Lifshitz. *Fluid Mechanics*. 2nd. Butterworth-Heinemann, 1987.

[39] G. Sigl. *Astroparticle Physics: Theory and Phenomenology*. 1st. Atlantis Press, 2017.

[40] R. Stephan and M. Brueggen. *Introduction to High-Energy Astrophysics*. 1st. Cambridge University Press, 2011.

[41] D. Caprioli. "Cosmic-ray acceleration in supernova remnants: non-linear theory revised". *Journal of Cosmology and Astrophysics* 2012, 038 (2012).

[42] T. K. Gaisser et al. "Gamma-Ray Production in Supernova Remnants". *The Astrophysical Journal* 492, 1 (1998).

[43] K. M. Schure et al. "Diffusive Shock Acceleration and Magnetic Field Amplification". *Space Science Reviews* 173, 1-4 (2012).

[44] O. Adriani et al. "Measurement of Boron and Carbon Fluxes in Cosmic Rays with the PAMELA Experiment". *The Astrophysical Journal* 791, 2 (2014).

[45] M. Aguilar et al. "Precision Measurement of the Boron to Carbon Flux Ratio in Cosmic Rays from 1.9 GV to 2.6 TV with the Alpha Magnetic Spectrometer on the International Space Station". *Physical Review Letters* 117, 231102 (2016).

[46] R. H. Kraichnan. "Inertial-range spectrum of hydromagnetic turbulence". *Physics of Fluids* 8, 7 (1965).

[47] A. N. Kolmogorov. "The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds Numbers". *Proceedings: Mathematical and Physical Sciences* 434, 1890 (1991).

[48] T. Stanev. *High Energy Cosmic Rays*. 2nd. Springer-Verlag Berlin Heidelberg, 2010.

[49] R. Aloisio. "Acceleration and propagation of ultra-high energy cosmic rays". *Progress of Theoretical and Experimental Physics* 2017, 12 (2017).

[50] M. Unger et al. "Highlights from the Pierre Auger Observatory (ICRC17)". *Proceedings of Science* PoS(ICRC2017)1102 (2017).

[51] A. Aab et al. "Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory". *Journal of Cosmology and Astroparticle Physics* 2017, 038 (2017).

[52] R. Abbasi et al. "Study of Ultra-High Energy Cosmic Ray composition using Telescope Array's Middle Drum detector and surface array in hybrid mode". *Astroparticle Physics* 64 (2015).

[53] M. Abe et al. "Depth of Ultra High Energy Cosmic Ray Induced Air Shower Maxima Measured by the Telescope Array Black Rock and Long Ridge FADC Fluorescence Detectors and Surface Array in Hybrid Mode". *Astrophysical Journal* 858, 76 (2018).

[54] S. Huemmer et al. "Energy dependent neutrino flavor ratios from cosmic accelerators on the Hillas plot". *Astroparticle Physics* 34, 4 (2010).

[55] C. Arguelles et al. "New Physics in Astrophysical Neutrino Flavor". *Physical Review Letters* 115, 161303 (2015).

[56] A. Esmaili and P. D. Serpico. "Are IceCube neutrinos unveiling PeV-scale decaying dark matter?" *Journal of Cosmology and Astroparticle Physics* 2013 (2013).

[57] D. S. Buhpal Dev et al. "Heavy right-handed neutrino dark matter and PeV neutrinos at IceCube". *Journal of Cosmology and Astroparticle Physics* 2016 (2016).

[58] M. Chianese et al. "Dark Matter interpretation of low energy IceCube MESE excess". *Journal of Cosmology and Astroparticle Physics* 2017, 007 (2017).

[59] C. Spiering. *Status and Perspectives of Astroparticle Physics in Europe*. Ed. by S. Roeser. Wiley VCH Verlag GmbH, 2008.

[60] R. Abbasi. "IceCube Sensitivity for Low-Energy Neutrinos from Nearby Supernovae". *Astronomy and Astrophysics* 535, A109 (2011).

[61] F. Vissani and F. Aharonian. "Galactic sources of high-energy neutrinos: Highlights". *Nuclear Instruments and Methods in Physics Research A* 692 (2012).

[62] M. G. Aartsen et al. "All-sky search for time-integrated neutrino emission from astrophysical sources with 7 years of IceCube data". *Astrophysical Journal* 835, 2 (2017).

[63] M. Ahlers. *Galactic Neutrino Sources*. Ed. by T. K. Gaisser and A. Karle. Vol. Neutrino Astronomy: Current Status, Future Prospects. World Scientific, 2017.

[64] B. P. Abbott et al. "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral". *Physical Review Letters* 119, 161101 (2017).

[65] E. Waxman and J. Bahcall. "High Energy Neutrinos from Cosmological Gamma-Ray Burst Fireballs". *Physical Review Letters* 78 (1997).

[66] P. Baerwald et al. "Are Gamma-Ray Bursts the Sources of Ultra-High Energy Cosmic Rays?" *Astroparticle Physics* 62 (2015).

[67] M. G. Aartsen et al. "An All-Sky Search for Three Flavors of Neutrinos from Gamma-Ray Bursts with the IceCube Neutrino Observatory". *Astrophysical Journal* 824, 2 (2016).

[68] M. Bustamante et al. "Neutrino and cosmic-ray emission from multiple internal shocks in gamma-ray bursts". *Nature Communications* 6, 6783 (2015).

[69] H. Hao-Ning et al. "Icecube Nondetection of Gamma-Ray Bursts: Constraints on the Fireball Properties". *The Astrophysical Journal* 752, 29 (2012).

[70] P. Meszaros. *Gamma-Ray Bursts as Neutrino Sources*. Ed. by T. K. Gaisser and A. Karle. Vol. Neutrino Astronomy: Current Status, Future Prospects. World Scientific, 2017.

[71] P. L. Biermann and P. A. Strittmatter. "Synchrotron emission from shock waves in active galactic nuclei". *Astrophysical Journal* 322 (1987).

[72] F. W. Stecker. "IceCube Observed PeV Neutrinos from AGN Cores". *Physical Review* D88, 047301 (2013).

[73] S. S. Kimura et al. "Neutrino and Cosmic-Ray Emission and Cumulative Background from Radiatively Inefficient Accretion Flows in Low-Luminosity Active Galactic Nuclei". *Astrophysical Journal* 806, 159 (2015).

[74] K. Fang and K. Murase. "Linking high-energy cosmic particles by black-hole jets embedded in large-scale structures". *Nature Physics* 14 (2018).

[75] E. Waxman. *The Origin of IceCube's Neutrinos: Cosmic Ray Accelerators Embedded in Star Forming Calorimeters*. Ed. by T. K. Gaisser and A. Karle. Vol. Neutrino Astronomy: Current Status and Future Prospects. World Scientific, 2017.

[76] A. Loeb and E. Waxman. "The cumulative background of high energy neutrinos from starburst galaxies". *Journal of Cosmology and Astroparticle Physics* 05, 003 (2006).

[77] I. Tamborra et al. "Star-forming galaxies as the origin of diffuse high-energy backgrounds: gamma-ray and neutrino connections, and implications for starburst history". *Journal of Cosmology and Astroparticle Physics* 9, 043 (2014).

[78] K. Bechtol et al. "Evidence against star-forming galaxies as the dominant source of IceCube neutrinos". *The Astrophysical Journal* 836, 1 (2017).

[79] K. Murase et al. "Hidden Cosmic-Ray Accelerators as an Origin of TeV-PeV Cosmic Neutrinos". *Physical Review Letters* 116, 071101 (2016).

[80] E. Middell. "Reconstruction of Cascade-Like Events in IceCube". MA thesis. Humboldt-Universitaet Berlin, 2008.

[81] E. Waxman and J. Bahcall. "High Energy Neutrinos from Astrophysical Sources: An Upper Bound". *Physical Review* D59, 023002 (1999).

[82] K. Mannheim et al. "On the cosmic ray bound for models of extragalactic neutrino production". *Physical Review* D63, 023003 (2001).

[83] E. Waxman and J. Bahcall. "High Energy Astrophysical Neutrinos: the Upper Bound is Robust". *Physical Review* D64, 023002 (2001).

[84] A. Neronov and D. V. Semikoz. "Neutrinos from Extra-Large Hadron Collider in the Milky Way". *Astroparticle Physics* 72, 32 (2015).

[85] M. Kachelriess. "Signatures of a two million year old supernova in the spectra of cosmic ray protons, antiprotons and positrons". *Physical Review Letters* 115, 181103 (2015).

[86] D. Gaggero et al. "The gamma-ray and neutrino sky: A consistent picture of Fermi-LAT, Milagro, and IceCube results". *The Astrophysical Journal Letters* 815, 25 (2015).

[87] M. G. Aartsen et al. "Constraints on Galactic Neutrino Emission with Seven Years of IceCube Data". *Astronomical Journal* 849, 67 (2017).

[88] R. Aloisio et al. "The UHECR source evolution and high-energy neutrinos and gamma-rays". *Proceedings of Science* PoS(ICRC2017)560 (2017).

[89] K. Greisen. "End of the Cosmic-Ray Spectrum". *Physical Review Letters* 16 (1966).

[90] G. T. Zatsepin and V. A. Kuzmin. "Upper Limit of the Spectrum of Cosmic Rays". *Journal of Experimental and Theoretical Physics Letters* 4 (1966).

[91] S. Yoshida. *Ultra-High Energy Particle Astrophysics*. Nova Science Pub Inc, 2003.

[92] R. Aloisio et al. "Ultra High Energy Cosmic Rays: The disappointing model". *Astroparticle Physics* 34, 620 (2011).

[93] M. Ahlers and F. Halzen. "Minimal Cosmogenic Neutrinos". *Physical Review* D86, 083010 (2012).

[94] M. G. Aartsen et al. "Constraints on Ultrahigh-Energy Cosmic-Ray Sources from a Search for Neutrinos above 10 PeV with IceCube". *Physical Review Letters* 117, 241101 (2016).

[95] P. S. Coppi and F. Aharonian. "Constraints on the VHE Emissivity of the Universe from the Diffuse GeV Gamma-Ray Background". *Astrophysical Journal* 487 (1997).

[96] V. S. Berezinsky et al. "Restricting UHECRs and cosmogenic neutrinos with Fermi-LAT". *Physics Letters B* 695 (2011).

[97] L. Raedel. "Measurement of High-Energy Muon Neutrinos with the IceCube Neutrino Observatory". PhD thesis. RWTH Aachen University, 2017.

[98] A. Fedynitch et al. "Calculation of conventional and prompt lepton fluxes at very high energy". *European Physical Journal Web of Conferences* 99, 08001 (2015).

[99] M. Honda et al. "Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data". *Physical Review* D75, 4 (2007).

[100] G. D. Barr et al. "A three-dimensional calculation of atmospheric neutrinos". *Physical Review* D70, 023006 (2004).

[101] R. Enberg et al. "Prompt neutrino fluxes from atmospheric charm". *Physical Review* D78, 043005 (2008).

[102] A. Bhattacharya et al. "Perturbative charm production and the prompt atmospheric neutrino flux in light of RHIC and LHC". *Journal of High Energy Physics* 06, 110 (2015).

[103] D. Chirkin. "Fluxes of Atmospheric Leptons at 600 GeV - 60 TeV". *ArXiv* hep-ph, 0407078 (2004).

[104] R. Gauld et al. "The prompt atmospheric neutrino flux in the light of LHCb". *Journal of High Energy Physics* 130 (2016).

[105] M. Benzke et al. "Prompt neutrinos from atmospheric charm in the general-mass variable-flavor-number scheme". *Journal of High Energy Physics* 12, 21 (2017).

[106] A. Bhattacharya et al. "Prompt atmospheric neutrino fluxes: perturbative QCD models and nuclear effects". *Journal of High Energy Physics* 11, 167 (2016).

[107] C. L. Cowan et al. "Detection of the Free Neutrino: a Confirmation". *Science* 124, 3212 (1956).

[108] D. G. Michael et al. "The magnetized steel and scintillator calorimeters of the MINOS experiment". *Nuclear Instruments and Methods in Physics Research A* 596 (2008).

[109] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos". *Physical Review Letters* 81 (1998).

[110] P. Gorham et al. "The Antarctic Impulsive Transient Antenna Ultra-high Energy Neutrino Detector Design, Performance, and Sensitivity for 2006-2007 Balloon Flight". *Astroparticle Physics* 32 (2009).

[111] F. Halzen and S. Klein. "IceCube: An Instrument for Neutrino Astronomy". *Review of Scientific Instruments* 81, 8 (2010).

[112] A. Cooper-Sarkar et al. "The high energy neutrino cross-section in the Standard Model and its uncertainty". *Journal of High Energy Physics* 8, 42 (2011).

[113] S. L. Glashow. "Resonant Scattering of Antineutrinos". *Physical Review* 118, 1 (1960).

[114] R. Gandhi et al. "Ultrahigh-Energy Neutrino Interactions". *Astroparticle Physics* 5 (1996).

[115] J. A. Formaggio and G. P. Zeller. "From eV to EeV: Neutrino Cross Sections Across Energy Scales". *Reviews of Modern Physics* 84 (2012).

[116] C. Patrignani et al. "Review of Particle Physics". *Chinese Physics C* 40, 100001 (2017).

[117] A. Cooper-Sarkar et al. "Proton Structure from HERA to LHC". *Proceedings of International Symposium on Mutliparticle Dynamics* (2010).

[118] H.-L. Lao et al. "New parton distributions for collider physics". *Physical Review* D82, 074024 (2010).

[119] I. E. Tamm and I. M. Frank. "Coherent Visible Radiation of Fast Electrons Passing Through Matter". *Doklady Akademii Nauk SSSR* 14, 109 (1937).

[120] N. J. Carron. *An Introduction to the Passage of Energetic Particles through Matter*. Taylor and Francis, 2007.

[121] C. Wiebusch and L. Raedel. "Calculation of the Cherenkov light yield from low energetic secondary particles accompanying high-energy muons in ice and water with Geant4 simulations". *Astroparticle Physics* 38 (2012).

[122] C. Wiebusch and L. Raedel. "Calculation of the Cherenkov-light yield from electromagnetic cascades in ice with Geant4". *Astroparticle Physics* 44, 102 (2013).

[123] D. Chirkin and W. Rhode. "Propagating leptons through matter with Muon Monte Carlo (MMC)". *arXiv* hep-ph, 0407075 (2004).

[124] IceCube Collaboration: R. Abbasi, et al. "Calibration and characterization of the IceCube photomultiplier tube". *Nuclear Instruments and Methods in Physics Research Section A* 618 (2010).

[125] IceCube Collaboration: and M. G. Aartsen. "Measurement of South Pole ice transparency with the IceCube LED calibration system". *Nuclear Instruments and Methods in Physics Research Section A* 711 (2013).

[126] M. G. Aartsen et al. "The IceCube Neutrino Observatory: Instrumentation and Online Systems". *Journal of Instrumentation* 12 (2017).

[127] R. Abbasi et al. "Calibration and Characterization of the IceCube Photomultiplier Tube". *Nuclear Instruments and Methods in Physics Research A* 618 (2010).

[128] D. Chirkin. "Evidence of optical anisotropy of the South Pole ice". *Proceedings of the 33rd ICRC* (2013).

[129] P. B. Price et al. "Age vs depth of glacial ice at South Pole". *Geophysical Research Letters* 27, 14 (2000).

[130] N. E. Bramall et al. "A deep high-resolution optical log of dust, ash and stratigraphy in South Pole glacial ice". *Geophysical Research Letters* 32, L21815 (2005).

[131] M. R. Legrand et al. "Vostok (Antarctica) ice core: Atmospheric chemistry changes over the last climatic cycle (160,000 years)". *Atmospheric Environment* 22 (1988).

[132] O. Watanabe et al. "The paleoclimate record in the ice core at Dome Fuji station, East Antarctica". *Annals of Glaciology* 29 (1999).

[133] M. Ackermann et al. "Optical properties of deep glacial ice at the South Pole". *Journal of Geophysical Research: Atmospheres* 111, D13 (2006).

[134] R. Abbasi et al. "The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping". *Nuclear Instruments and Methods in Physics Research A* 601 (2009).

[135] M. Lesiak-Bzdak and L. Schulte. "Request for the Online Cascade Filter and Additional Calculations at Level 2 for the IC86 2013 Cascade Filter Stream". *IceCube Internal Communications* (2012).

[136] M. Lesiak-Bzdak and C. H. Ha. "Level3 Cuts for IC86 Cascade Channel". *IceCube Internal Communications* (2013).

[137] IceCube Collaboration: M. G. Aartsen, et al. "Energy Reconstruction Methods in the IceCube Neutrino Telescope". *Journal of Instrumentation* 9, P03009 (2014).

[138] R. Abbasi. "Search for ultrahigh-energy tau neutrinos with IceCube". *Physical Review* D86, 022005 (2012).

[139] J. van Santen. "Neutrino Interactions in IceCube above 1 TeV". PhD thesis. University of Wisconsin-Madison, 2014.

[140] J. Lundberg et al. "Light tracking through ice and water – Scattering and absorption in heterogeneous media with Photonics". *Nuclear Instruments and Methods in Physics Research A* A581 (2007).

[141] M. Kowalski. "Search for Neutrino-Induced Cascades with the AMANDA-II Detector". PhD thesis. Humboldt-Universitaet Berlin, 2003.

[142] R. Abbasi et al. "Search for neutrino-induced cascades with five years of AMANDA data". *Astroparticle Physics* 34 (2011).

[143] D. Pandel. "Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal-Neutrinoteleskop NT-72". MA thesis. Humboldt-Universitaet Berlin, 1996.

[144] J. van Santen. "Markov-Chain Monte-Carlo Reconstruction for cascade-like events in IceCube". MA thesis. Humboldt-Universitaet Berlin, 2010.

[145] J. Ahrens et al. "Muon Track Reconstruction and Data Selection Techniques in AMANDA". *Nuclear Instruments and Methods in Physics Research A* 524 (2004).

[146] R. J. Lauer. "Extending the search for cosmic point sources of neutrinos with IceCube beyond PeV energies and above the horizon". PhD thesis. Humboldt-Universitaet Berlin, 2010.

[147] A. Gazizov and M. P. Kowalski. "ANIS: High Energy Neutrino Generator for Neutrino Telescopes". *Computer Physics Communications* 172 (2005).

[148] J. H. Koehne. "PROPOSAL: A tool for propagation of charged leptons". *Computer Physics Communications* 184 (2013).

[149] C. Kopper. "CLsim Software". *IceCube Internal Software Repository* (2018).

[150] D. Chirkin. "PPC Software". *IceCube Internal Software Repository* (2018).

[151] D. Heck et al. *CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers*. Tech. rep. FZKA 6019. Forschungszentrum und Universität Karlsruhe, 1998.

[152] T. Chen and C. Guestrin. "XGBoost: A Scalable Tree Boosting System". *Proceedings of 22nd KDD Conference* (2016).

[153] Y. Xu. "Measurement of the TeV-PeV Neutrino-Nucleon Cross Section with Five Years of IceCube Data". PhD thesis, Stony Brook University, in prep. 2018.

[154] A. Stoessl. "A search for particle showers at the edge of IceCube's instrumental volume". PhD thesis. Humboldt-Universitaet Berlin, 2015.

[155] IceCube Collaboration: M. G. Aartsen, et al. "High Energy Astrophysical Neutrino Flux Characteristics for Neutrino-induced Cascades Using IC79 and IC86-String IceCube Configurations". *Proceedings of Science* PoS(ICRC2015)1109 (2015).

[156] L. Lu. "Multi-flavour PeV neutrino search with IceCube". *Proceedings of Science* PoS(ICRC2017)1002 (2017).

[157] J. Friedman. "Greedy Function Approximation: A Gradient Boosting Machine". *The Annals of Statistics* 29, 5 (2001).

[158] J. Friedman et al. "Additive logistic regression: a statistical view of boosting". *The Annals of Statistics* 28, 2 (2000).

[159] T. Hastie et al. *The Elements of Statistical Learning*. 1st. Springer-Verlag Berlin Heidelberg, 2001.

[160] C. Bishop. *Pattern Recognition and Machine Learning*. Springer-Verlag Berlin Heidelberg, 2011.

[161] C. Adam-Bourdarios et al. "The Higgs boson machine learning challenge". *MLR: Workshop and Conference Proceedings 42* (2015).

[162] D. Nielsen. "Tree Boosting With XGBoost - Why Does XGBoost Win "Every" Machine Learning Competition?" MA thesis. Norwegian University of Science and Technology, 2016.

[163] R. Kohavi. "A study of cross-validation and bootstrap for accuracy estimation and model selection". *Proceedings of the 14th IJCAI conference* (1995).

[164] IceCube Collaboration: M. G. Aartsen, et al. "Atmospheric and Astrophysical Neutrinos above 1 TeV Interacting in IceCube". *Physical Review* D91, 2 (2015).

[165] IceCube Collaboration: M. G. Aartsen, et al. "The IceCube Neutrino Observatory: Instrumentation and Online Systems". *Journal of Instrumentation* 12, P03012 (2017).

[166] IceCube Collaboration: M. Rongen, et al. "Measuring the optical properties of IceCube drill holes". *EPJ Web of Conferences* 116 (2016).

[167] A. Fedynitch et al. "Influence of hadronic interaction models and the cosmic-ray spectrum on the high-energy atmospheric muon and neutrino flux". *Physical Review* D86, 11 (2012).

[168] T. Pierog. "Air Shower Simulation with a New Generation of post-LHC Hadronic Interaction Models in CORSIKA". *Proceedings of Science* PoS(ICRC2017)1100 (2017).

[169] A. Fedynitch et al. "A state-of-the-art calculation of atmospheric lepton fluxes". *Proceedings of Science* PoS(ICRC2017)1019 (2017).

[170] T. K. Gaisser. "Spectrum of cosmic-ray nucleons and the atmospheric muon charge ratio". *Astroparticle Physics* 35, 12 (2012).

[171] D. Heck and T. Pierog. *Extensive Air Shower Simulation with CORSIKA: A User's Guide*. Tech. rep. Forschungszentrum und Universität Karlsruhe, 2011.

[172] NASA. *U.S. Standard Atmsophere, 1976*. Tech. rep. NASA-TM-X-74335. National Aeronautics and Space Administration, 1976.

[173] F. Riehn et al. "A new version of the event generator Sibyll". *Proceedings of Science* PoS(ICRC2015)558 (2015).

[174] S. Ostapchenko. "Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model". *Physical Review* D83, 1 (2011).

[175] T. Pierog et al. "EPOS LHC : test of collective hadronization with LHC data". *Physical Review* C92 (2015).

[176] S. Roesler et al. "The Monte Carlo Event Generator DPMJET-III". Springer-Verlag Berlin, 2001. Chap. The Monte Carlo Event Generator DPMJET-III.

[177] A. Fedynitch and R. Engel. "Revision of the high energy hadronic interacion models PHOJET/DPMJETIII". *Proceedings of 14th International Conference on Nuclear Reaction Mechanisms* (2015).

[178] Y. S. Jeong. "Prompt atmospheric neutrino flux in perturbative QCD and its theoretical uncertainties". *Journal of Physics: Conference Series* 888, 012117 (2017).

[179] J. van Santen et al. "Generalized self-veto probability for atmospheric neutrinos". *Physical Review* D90, 2 (2014).

[180] S. Schönert et al. "Vetoing atmospheric neutrinos in a high energy neutrino telescope". *Physical Review* D79, 4 (2009).

[181] IceCube Collaboration: M. G. Aartsen, et al. "Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data". *Physical Review Letters* 113, 10 (2014).

[182] IceCube Collaboration: M. G. Aartsen, et al. "Observation of Astrophysical Neutrinos in Four Years of IceCube Data". *Proceedings of Science* PoS(ICRC2015)1081 (2015).

[183] G. McLachlan and D. Peel. *Finite Mixture Models*. 1st. Wiley Series in Probability and Statistics. Wiley-Interscience, 2000.

[184] G. Ambrosi et al. "Direct detection of a break in the teraelectronvolt cosmic-ray spectrum of electrons and positrons". *nature* 552 (2017).

[185] E. Massaro et al. "Log-parabolic spectra and particle acceleration in blazars". *Astronomy and Astrophysics* 448 (2006).

[186] C. D. Dermer et al. "Photopion production in black-hole jets and flat-spectrum radio quasars as PeV neutrino sources". *Journal of High Energy Physics* 3-4 (2014).

[187] L. Wasserman. *All of Nonparametric Statistics*. Springer, 2006.

[188] J. M. Picone et al. "NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues". *Journal of Geophysical Research: Space Physics* 107, A12 (2002).

[189] G. Casella and R. L. Berger. *Statistical Inference*. 2nd. Duxbury, 2002.

[190] J. O. Berger and R. L. Wolpert. *The Likelihood Principle*. 2nd. Hayward, CA: Institute of Mathematical Statistics, 1988.

[191] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". *The Annals of Mathematical Statistics* 9, 1 (1938).

[192] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". *The European Physical Journal* C71 (2011).

[193] H. Chernoff. "On the Distribution of the Likelihood Ratio". *The Annals of Mathematical Statistics* 25, 3 (1954).

[194] L. Demortier. *P-Values: What They Are and How to Use Them*. Tech. rep. 8662. CDF/MEMO/STATISTICS/PUBLIC/, 2007.

[195] O. Behnke et al. *Data Analysis in High Energy Physics*. 1st. WILEY-VCH, 2013.

[196] G. J. Feldman and R. D. Cousins. "Unified approach to the classical statistical analysis of small signals". *Physical Review* D57, 7 (1998).

[197] B. Sen et al. "On the Unified Method with Nuisance Parameters". *Statistica Sinica* 19 (2009).

[198] S. Baker and R. D. Cousins. "Clarification of the Use of Chi-square and Likelihood Functions in Fits to Histograms". *Nuclear Instruments and Methods in Physics Research* 221 (1984).

[199] IceCube Collaboration: M. G. Aartsen, et al. "A Combined Maximum-likelihood Analysis of the High-energy Astrophysical Neutrino Flux Measured with IceCube". *The Astrophysical Journal* 809, 98 (2015).

[200] A. Schukraft. "Search for a diffuse flux of astrophysical muon neutrinos with the IceCube 59-string configuration". *Physical Review* D89, 062007 (2014).

[201] IceCube Collaboration: M. G. Aartsen, et al. "Observation and Characterization of a Cosmic Muon Neutrino Flux From the Northern Hemisphere Using Six Years of Icecube Data". *The Astrophysical Journal* 833, 1 (2016).

[202] IceCube Collaboration: M. G. Aartsen, et al. "All-flavor Multi-Channel Analysis of the Astrophysical Neutrino Spectrum with IceCube". *Proceedings of Science* PoS(ICRC2017)976 (2017).

[203] IceCube Collaboration: and M. G. Aartsen. "Measurement of Atmospheric Neutrino Oscillations at 6-56 GeV with IceCube DeepCore". *Physical Review Letters*, 120 (2018).

[204] IceCube Collaboration: M. G. Aartsen, et al. "Observation of Astrophysical Neutrinos in Six Years of IceCube Data". *Proceedings of Science* PoS(ICRC2017)981 (2017).

[205] IceCube Collaboration: M. G. Aartsen, et al. "A Measurement of the Diffuse Astrophysical Muon Neutrino Flux Using Eight Years of IceCube Data". *Proceedings of Science* PoS(ICRC2017)1005 (2017).

[206] M. G. Aartsen. "Search for a diffuse flux of astrophysical muon neutrinos with the IceCube 59-string configuration". *Physical Review* 2014, 062007 (D89).

[207] M. G. Aartsen. "Evidence for Astrophysical Muon Neutrinos from the Northern Sky with IceCube". *Physical Review Letters* 115, 081102 (2015).

[208] R. Abbasi et al. "First search for atmospheric and extraterrestrial neutrino-induced cascades with the IceCube detector". *Physical Review* D84, 072001 (2011).

[209] M. G. Aartsen. "Search for neutrino-induced particle showers with IceCube-40". *Physical Review* D89, 102001 (2014).

[210] H. Niederhausen and S. Schoenen. *Comparing Independent Multivariate Measurements: The General Case*. Internal Reports. IceCube Collaboration, 2015.

[211] R. B. Davies. "Hypothesis testing when a nuisance parameter is present only under the alternative". *Biometrika* 74, 33-43 (1987).

[212] E. Gross and O. Vitells. "Trial factors for the look elsewhere effect in high energy physics". *The European Physical Journal C* 70, 1-2 (2010).

[213] S. Algeri et al. "On methods for correcting for the look-elsewhere effect in searches for new physics". *Journal of Instrumentation* 11, P12010 (2016).

[214] S. Algeri and D. A. van Dyk. "Testing One Hypothesis Multiple Times". *arXiv [stat.ME]* 1701.06820 (2017).

[215] S. Schoenen. "Discovery and Characterization of a Diffuse Astrophysical Muon Neutrino Flux with the IceCube Neutrino Observatory". PhD thesis. RWTH Aachen University, 2017.

[216] S. D. Biller and S. M. Oser. "Another look at confidence intervals: Proposal for a more relevant and transparent approach". *Nuclear Instruments and Methods in Physics Research Section A* 774 (2015).

[217] J. P. Shaffer. "Multiple Hypothesis Testing". *Annual Review of Psychology* 46 (1995).

[218] A. Palladino and F. Vissani. "Extragalactic Plus Galactic Model For IceCube Neutrino Events". *The Astrophysical Journal* 826, 2 (2016).

[219] A. Palladino and W. Winter. "A Multi-Component Model for the Observed Astrophysical Neutrinos". *arXiv:1801.07277* (2018).

[220] A. C. Vincent et al. "Analysis of the 4-year IceCube high-energy starting events". *Physical Review* D94, 2 (2016).

[221] A. Gelman et al. *Bayesian Data Analysis*. 3rd. CRC Press, 2014.

[222] A. Palladino. "How large is the galactic contribution to IceCube neutrino events?" *Journal of Physics: Conference Series* 888, 1 (2017).

[223] A. Neronov and D. V. Semikoz. "Neutrinos from Extra-Large Hadron Collider in the Milky Way". *Astroparticle Physics* 72 (22016).

[224] M. G. Aartsen et al. "Search for astrophysical sources of neutrinos using cascade events in IceCube". *Astrophysical Journal* 846, 136 (2017).

[225] M. Ackermann et al. "The spectrum of isotropic diffuse gamma-ray emission between 100 MeV and 820 GeV". *Astrophysical Journal* 799, 86 (2015).

[226] F. Fenu et al. "The cosmic ray energy spectrum measured using the Pierre Auger Observatory". *Proceedings of Science* PoS(ICRC2017)486 (2017).

[227] S. Yoshida et al. "Differential limit on an EHE neutrino flux component in the presence of astrophysical background from nine years of IceCube data". *Proceedings of Science* PoS(ICRC2017)975 (2017).

[228] IceCube Collaboration: M. G. Aartsen, et al. "Seasonal variation of atmospheric neutrinos in IceCube". *Proceedings of the 33rd ICRC* (2013).