JOHANNES GUTENBERG UNIVERSITÄT MAINZ

JG|U

Stacked searches for high-energy neutrinos from blazars with IceCube

Kai Schatto
geboren in Bad Kreuznach

Fachbereich Physik, Mathematik und Informatik

Johannes Gutenberg-Universität Mainz

Dissertation zur Erlangung des Grades

„Doktor der Naturwissenschaften“

Mainz, 2. Juni 2014

1. Gutachter: aus datenschutzrechtlichen Gründen entfernt

2. Gutachter: aus datenschutzrechtlichen Gründen entfernt

Datum der mündlichen Doktorprüfung: 2. Oktober 2014

Kai Schatto
ETAP
Institut für Physik
Staudingerweg 7
Johannes Gutenberg-Universität Mainz
D-55128 Mainz
schattok@uni-mainz.de

Abstract

In the year 2013, the detection of a diffuse astrophysical neutrino flux with the IceCube neutrino telescope – constructed at the geographic South Pole – was announced by the IceCube collaboration. However, the origin of these neutrinos is still unknown as no sources have been identified to this day. Promising neutrino source candidates are blazars, which are a subclass of active galactic nuclei with radio jets pointing towards the Earth.

In this thesis, the potential neutrino flux from blazars is tested with a maximum likelihood stacking approach, analyzing the combined emission from uniform groups of objects. The stacking enhances the sensitivity with respect to the as-yet unsuccessful single source searches. The analysis utilizes four years of IceCube data including one year from the completed detector. As all results presented in this work are compatible with background, upper limits on the neutrino flux are given. It is shown that, under certain conditions, some hadronic blazar models can be challenged or even rejected.

Moreover, the sensitivity of this analysis – and any other future IceCube point source search – was enhanced by the development of a new angular reconstruction method. It is based on a detailed simulation of the photon propagation in the Antarctic ice. The median resolution for muon tracks induced by high-energy neutrinos is improved for all neutrino energies above IceCube's lower threshold of $\approx 0.1\,\mathrm{TeV}$. By reprocessing the detector data and simulation from the year 2010, it is shown that the new method improves IceCube's discovery potential by 20 % to 30 % depending on the declination.

Contents

List of Figures ix

List of Tables xiii

1 Introduction 1

2 Theory and detector 3
2.1 Neutrinos in and beyond the Standard Model of particle physics 3
2.2 Cosmic rays and neutrino astronomy 5
2.3 Possible sources of astrophysical neutrinos 8
2.3.1 Active galactic nuclei 10
2.3.1.1 Blazar emission models 13
2.4 High-energy neutrino detection 15
2.4.1 Muon propagation 17
2.4.2 The Cherenkov effect 18
2.4.3 Photon propagation in the deep Antarctic ice 19
2.4.4 Cosmic ray backgrounds 20
2.5 The IceCube neutrino telescope 22
2.5.1 The digital optical module (DOM) 23
2.5.2 Calibration and data quality 25
2.5.3 Data acquisition 26
2.5.4 Optical properties of the Antarctic ice 27
2.5.5 Simulation 28
2.5.5.1 IceTray - The IceCube software framework 28
2.5.5.2 Generators 29
2.5.5.3 Muon propagation 29
2.5.5.4 Photon propagation 30
2.5.5.5 Detector simulation 31
2.5.6 Recent results 31

3 Improved muon track reconstruction 35
3.1 Muon track reconstruction 35

3.1.1 Line-fit . . . 36
3.1.2 Likelihood reconstructions . . . 37
3.1.2.1 The Pandel function . . . 39
3.1.3 Bootstrapping . . . 40
3.1.4 Summary of the discussion . . . 41
3.2 Light table production: towards a new PDF . . . 42
3.3 Table-fitting with Photospline . . . 45
3.3.1 Photospline . . . 46
3.3.2 Fitting procedure . . . 47
3.4 Spline reconstruction results . . . 50
3.5 Further likelihood-fit improvements . . . 51
3.5.1 Hit cleaning with the Kolmogorov-Smirnov test . . . 52
3.5.2 Including mean stochastic energy losses . . . 54
3.5.3 Accurate noise modeling . . . 56
3.5.4 MPE likelihood convolution . . . 58
3.5.5 Combination . . . 60
3.6 Summary . . . 63

4 Data filtering and event selection 65
4.1 Cut and quality parameters for the analysis . . . 65
4.1.1 Energy estimators . . . 66
4.1.1.1 Analytical implementation . . . 67
4.1.1.2 Improved analytical implementation . . . 68
4.1.1.3 Truncated Energy . . . 69
4.1.2 Angular resolution estimation . . . 69
4.1.3 IceTop veto . . . 70
4.1.4 Hit topology variables . . . 71
4.1.5 Hit-track topology variables . . . 71
4.1.6 Reconstruction quality variables . . . 72
4.1.7 Reconstruction comparisons . . . 72
4.1.8 Status flags . . . 73
4.2 Basic data processing . . . 73
4.2.1 L1 and L2: filtering and first reconstructions . . . 73
4.2.2 L3: advanced routines and cuts . . . 74
4.2.3 L3b: Adding the SplineMPE reconstruction . . . 75
4.3 Multivariate classification with boosted decision trees . . . 76
4.3.1 BDT setup . . . 78
4.3.2 Input variables . . . 80
4.3.3 BDT configuration . . . 81
4.3.4 Pre-classification cuts . . . 83
4.3.5 BDT cut optimization . . . 84
4.3.6 Adjustment of the downgoing event rate . . . 86
4.4 Final sample properties and summary . . . 87
4.4.1 Energy estimator choice . . . 88

4.4.2 Paraboloid pull correction . . . 89
4.4.3 Angular resolution and sensitivity . . . 89

5 Analysis: Neutrinos from blazars 93
5.1 Which blazars are neutrino loud? . . . 93
5.1.1 Model uncertainties . . . 96
5.2 Analysis method . . . 96
5.2.1 Discovery potential, sensitivity and limits . . . 98
5.3 Source weighting . . . 99
5.4 Blazar selection . . . 99
5.5 Dominant sources and cutoff study . . . 100
5.6 Analysis data . . . 101
5.7 Expected performance . . . 102
5.8 Systematic uncertainties . . . 103
5.9 Results of the blazar stacking . . . 104
5.10 Model specific limits . . . 106

6 Summary 109

Appendices 111

A Data quality monitoring 111

B Additional tables and figures 115

References 127

List of Figures

2.1 Cosmic ray spectrum . . . 6
2.2 Hillas criterion . . . 8
2.3 Unified AGN model . . . 12
2.4 Optical spectra of FSRQs and BL Lacs . . . 13
2.5 Spectral energy distribution of 3C 66A . . . 14
2.6 Cherenkov light signatures . . . 15
2.7 Neutrino cross sections and interaction lengths . . . 17
2.8 Muon energy loss per meter . . . 18
2.9 Average muon range in ice . . . 18
2.10 Vertical atmospheric muon flux . . . 21
2.11 Average northern hemisphere atmospheric muon neutrino flux . . . 21
2.12 Signal types in IceCube . . . 22
2.13 The IceCube detector . . . 23
2.14 The digital optical module . . . 24
2.15 DOM electronics block diagram . . . 24
2.16 Waveform hit extraction . . . 26
2.17 Spice Mie ice model . . . 28
2.18 High-energy starting event analysis . . . 32

3.1 Muon track parameters . . . 36
3.2 The Pandel function . . . 40
3.3 Angular reconstruction comparison I . . . 40
3.4 Photon propagation coordinate system . . . 43
3.5 B-splines of order 0, 1 and 2 . . . 46
3.6 CDF spline fit . . . 49
3.7 PDF spline fit . . . 49
3.8 Expected light yield fit . . . 49
3.9 Spline convolution . . . 49
3.10 Angular reconstruction comparison II (with SplineMPE) . . . 50
3.11 Point spread function representations . . . 51
3.12 Expected photo electrons from muon with cascade . . . 52
3.13 Kolmogorov-Smirnov statistic . . . 52

3.14 Reconstruction with Kolmogorov-Smirnov cleaned pulses 54
3.15 Effective angular Cherenkov light distribution 55
3.16 PDFs of bare muon and averaged secondaries 55
3.17 Expected photo electrons depending on muon energy 56
3.18 Reconstruction results including averaged stochastic energy losses 56
3.19 Distribution of uncleaned noise hits 57
3.20 Distribution of cleaned noise hits 57
3.21 Number of hits per hit DOM for different hit types 58
3.22 Exact noise model reconstruction results 58
3.23 Convolved MPE likelihood 59
3.24 MPE Gaussian convolution results 60
3.25 Combined SplineMPE likelihood modifications 62

4.1 Energy resolution of MuE, MuEX and TE 68
4.2 Muon energy estimation for atmospheric muons and E^{-2} neutrinos from TE and MuE 68
4.3 Support points and Paraboloid fit 69
4.4 Correlation of Paraboloid and true deviation 69
4.5 IceTop veto efficiency 70
4.6 L3b data flow 76
4.7 BDT structure 77
4.8 BDT Results for the hard spectrum BDT in the horizontal region 85
4.9 Discovery potential vs. BDT cut value 85
4.10 Zenith distribution of the final sample 86
4.11 E^{-2} signal efficiency w.r.t L3 86
4.12 Energy-zenith distribution of the final sample 87
4.13 IC79b final sample effective area 88
4.14 Sensitivity comparison between MuEX and MuE 88
4.15 Uncorrected and corrected SplineMPE Paraboloid pull 89
4.16 SplineMPE angular resolution in IC79b 91
4.17 E^{-2} sensitivity and discovery potential of the IC79b sample 91

5.1 γ-attenuation on the extragalactic background light (EBL) 94
5.2 Integrated flux over 3C 279 flare 94
5.3 Mkn 421 and PKS 0716+714 flux prediction 95
5.4 PG 1553 +11 flux prediction 95
5.5 P-value evaluation 97
5.6 Energy PDF vs. reconstructed energy 97
5.7 Generic stacking limits 105
5.8 Limits on specific flux models 107

A.1 Monitoring run summary 111
A.2 Monitoring trigger/filter rates 112
A.3 Monitoring DOM map 113

A.4 Reported data quality and detector issues 114

B.1 Reconstruction with Kolmogorov-Smirnov cleaned Pulses for E^{-3} 123
B.2 MPE Gaussian convolution results for an E^{-3} neutrino spectrum 123
B.3 E^{-1} sensitivity and discovery potential of the IC79b sample 123
B.4 E^{-3} sensitivity and discovery potential of the IC79b sample 123
B.5 Examples for signal-background separating variables on L3b 124
B.6 Data-Monte Carlo comparison 125

List of Tables

2.1 AGN types and emission properties 11

3.1 Photon table binning 44
3.2 Photospline knot configuration 47
3.3 Uncertainties in the IceCube time measurement 59

4.1 Reconstructions on level 3 75
4.2 IC79b BDT setup 79
4.3 Input variables of the upgoing BDTs 80
4.4 Input variables of the downgoing BDTs 81
4.5 Performance of different BDT configurations for the downgoing BDT1 . . . 82
4.6 BDT settings 82
4.7 Pre-cut efficiency 84
4.8 BDT cut values 85

5.1 Blazar catalog selection criteria 100
5.2 Dominant sources 101
5.3 Analysis data statistics 102
5.4 Blazar stacking sensitivity and discovery potential 103
5.5 Systematic uncertainties 104
5.6 Blazar stacking results 105

B.1 Critical values for the Kolmogorov-Smirnov test 116
B.2 Performance of different BDT configurations for the downgoing BDT2 . . . 117
B.3 L3b software setup 117
B.4 Adjusted BDT cut values in the downgoing region 118
B.5 BDT cut values in the overlap zone 119
B.6 FSRQ catalog 120
B.7 LSP BL Lac catalog 121
B.8 Hard γ-spectrum BL Lac catalog 122

1 Introduction

In the last century, tremendous progress was made in many fields of physics. The Standard Model of particle physics very successfully describes the interactions of fundamental particles and cosmology is constantly clearing the fog around the evolution of the universe. However, while we researched nature simultaneously on subatomic and astronomical scales, an astonishing fact became more and more clear: The largest phenomena in the universe are connected to the smallest processes of elementary particles to a much larger extent than we ever anticipated. We are not able to explain the enormous explosions of massive stars without knowledge of the most elusive particles – the neutrinos. We are not able to understand the vast electromagnetic emission from the most powerful persistent objects in the universe, active galactic nuclei, without research on fundamental leptonic and hadronic interactions and the detection or non-detection of neutrinos. One logical consequence from this insight is the field of astroparticle physics and especially the neutrino astronomy with the currently largest neutrino telescope IceCube.

While the IceCube collaboration announced the first detection of a diffuse astrophysical neutrino flux in the year 2013, specific high-energy neutrino sources could not yet be identified. Thus, while data is accumulating continuously, the analysis methods have to be improved in parallel. In the course of this work, a new angular reconstruction method is developed, significantly improving the resolution for neutrino induced high-energy muons. The method is based on a detailed simulation of the photon propagation in IceCube's detection medium – the Antarctic ice sheet. By reprocessing the IceCube data from the year 2010, the improvement in discovery potential with respect to older methods will be evaluated.

Furthermore, maximum likelihood stacking analyses improve the sensitivity with respect to single source searches by analyzing the flux sum from a uniform group of objects. In this work, the method will be used to investigate the neutrino flux from a subclass of active galactic nuclei with radio jets pointing towards the Earth – so called blazars. Three

promising blazar groups will be identified from theoretical arguments and analyzed with the stacking method, utilizing four years of IceCube data.

The thesis starts with the required theoretical principles and a description of the IceCube detector in Chapter 2. In Chapter 3, the development of the improved muon reconstruction method is described. Afterwards, Chapter 4 discusses the application of this new reconstruction to the data of the year 2010 which is used in the blazar stacking analysis covered in Chapter 5. Finally, the results are summarized in Chapter 6.

As can be seen from this short outline – touching on topics like enormous blazars, single photon propagation and cubic kilometer sized detectors – the connection of the largest and tiniest scales mentioned above is also a leitmotif for this thesis.

2

Theory and detector

The introduction provides the basic concepts necessary for the understanding of this work. Section 2.1 briefly reviews the role of the neutrino in and beyond the Standard Model of particle physics. Afterwards, in Section 2.2, I describe the connection between cosmic rays and astrophysical neutrinos and the motivation for neutrino astronomy. Section 2.3 discusses potential sources of astrophysical neutrinos, and Section 2.4 introduces the main principles of high-energy neutrino detection. Finally, I provide a detailed description of the IceCube detector in Section 2.5.

2.1 Neutrinos in and beyond the Standard Model of particle physics

The Standard Model of particle physics, developed in the 20th century, describes the properties and interactions of all known subatomic particles. It is a quantum field theory defined by the gauge symmetry group $SU(3) \times SU(2) \times U(1)$ describing three elementary forces, namely the strong, weak and electromagnetic interaction. The latter two have been unified in the electroweak force, which is represented by the symmetry group $U(1) \times SU(2)_L$. These fundamental interactions are conveyed by spin 1 particles called gauge bosons. The electromagnetic force is transmitted by photons, the strong coupling by 8 gluons, and the weak interaction by Z^0 and $W^{\pm}$ bosons.

The group of fermions is formed by 12 particles with spin 1/2. Hadronic matter consists of combinations from six quarks (up, down, charm, strange, top, bottom), which are arranged in three generations. A lepton generation is formed by a charged lepton (e, μ, τ) and the correspondent uncharged neutrino (ν_e, ν_μ, ν_τ). Neutrinos only interact weakly and

have zero mass in the Standard Model. However, the observation of neutrino oscillations[1] indicated that flavor eigenstates are a superposition of different mass eigenstates, and thus neutrinos carry a very small nonzero mass[2].

The flavor eigenstates $|\nu_a\rangle$, with $a = e, \mu, \tau$, at time t are given by the mass eigenstates $|\nu_j\rangle$, with $j = 1, 2, 3$, according to [4, 5]

$$|\nu_a(t)\rangle = \sum_{j=1,2,3} U_{aj} \exp(-iE_{\nu,j}t) \, |\nu_j\rangle \ , \tag{2.1}$$

with the mass eigenstate energy $E_{\nu,j}$ and $\hbar = c = 1$. The lepton mixing matrix U (also called PMNS matrix[3]) is given by [4, 5]

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} , \tag{2.2}$$

where s and c indicate the sine and cosine of the three mixing angles θ_{12}, θ_{13}, θ_{23}, and δ denotes the CP violating phase [8]. The probability that a neutrino, produced with flavor a, oscillates to a flavor b is then given by [5]

$$P_{a\to b} = |\langle \nu_b(t) | \nu_a(t=0) \rangle|^2 = \delta_{ab} - 4 \sum_{j>i} U^*_{ai} U_{bi} U_{aj} U^*_{bj} \sin^2 \left(\frac{\Delta m^2_{ij} L}{4E_\nu} \right) , \tag{2.3}$$

with the traveled distance L and the difference of the squared masses $\Delta m^2_{ij} = m^2_i - m^2_j$. Of special interest for neutrino astronomy is the fact that for large[4] L, the $\sin^2$ expression averages to ≈ 0.5 [9] and Equation 2.3 simplifies to $\sum_j |U_{aj}|^2 |U_{bj}|^2$. This means, a flavor composition constant in time and distance is expected from astrophysical neutrinos that only depends on the flavor composition at the production site. With recent experimental values for the mixing angles and δ [10]

$$\theta_{12} = 0.59 \, , \tag{2.4}$$
$$\theta_{13} = 0.16 \, , \tag{2.5}$$
$$\theta_{23} = 0.67 \ \text{and} \tag{2.6}$$
$$\delta = 1.08 \cdot \pi \, , \tag{2.7}$$

the oscillated flavor composition can be given from Equation 2.3 as [11]

[1] First indications for neutrino oscillation came from the Homestake experiment measuring the solar neutrino deficit [1]. It was shown at the Sudbury Neutrino Observatory and various other experiments that the deficit is caused by flavor oscillations [2].

[2] The current upper limit for the ν_e mass is $< 2\,\text{eV}$ [3].

[3] Named after the authors first describing the concept: Pontecorvo [6], Maki, Nakagawa and Sakata [7].

[4] Larger than the solar system [5].

$$\begin{pmatrix} \nu_e^{\text{Detected}} \\ \nu_\mu^{\text{Detected}} \\ \nu_\tau^{\text{Detected}} \end{pmatrix} = \begin{pmatrix} 0.548 & 0.244 & 0.208 \\ & 0.404 & 0.352 \\ & & 0.439 \end{pmatrix} \cdot \begin{pmatrix} \nu_e^{\text{Source}} \\ \nu_\mu^{\text{Source}} \\ \nu_\tau^{\text{Source}} \end{pmatrix} . \tag{2.8}$$

It will be shown later, that a production flavor ratio of $(\nu_e : \nu_\mu : \nu_\tau) \approx (1 : 2 : 0)$ is expected at astrophysical neutrino sources from common neutrino production scenarios via pion decays. In this case, Equation 2.8 delivers a ratio of $\approx (1 : 1 : 1)$ at earth. Thus, the flavor composition allows for direct conclusions on the physics processes at astrophysical sources.

Regarding the $\sin^2$ term in Equation 2.3, it is apparent that vacuum oscillation observations are not sensitive to the sign of Δm_{ij}^2 in first order. However, the mass hierarchy of the neutrino mass eigenstates can be determined by matter effects on the oscillation (MSW effect [12]). From oscillation effects in the solar medium it was determined that $\Delta m_{21}^2 = +7.5\,\text{eV}$ [13] while the sign of Δm_{32}^2 is still unknown. Thus, the mass hierarchy is one of the open questions in the neutrino sector. Moreover, it is unknown whether the neutrino is its own antiparticle (Majorana particle) and if the oscillation violates CP symmetry ($\delta \neq 0$). However, these issues are of subordinate relevance for this work and will therefore not be discussed in detail. Instead, Section 2.2 focuses on the importance of the neutrino as a cosmic messenger and the motivation for neutrino astronomy.

2.2 Cosmic rays and neutrino astronomy

Even 100 years after the discovery of cosmic rays in balloon experiments by Viktor Hess in 1912 [14], the origin and exact production mechanisms of these charged particles are not fully understood. However, the spectrum of cosmic rays has been measured to great detail by various experiments, shown in Figure 2.1. Up to energies around 100 TeV, the flux still allows for a direct measurement in balloon or satellite experiments, which are capable to identify the particle type [15]. At 10 GeV, the spectrum consists of 94 % protons, 5 % helium nuclei and 1 % heavier nuclei [16]. The steeply falling spectrum and corresponding low rates require indirect measurements by surface arrays at higher energies, monitoring the large particle showers created by cosmic ray interactions in the atmosphere. In this case, the composition is very tricky to measure [15].

The total spectrum spans over ten orders of magnitude, reaching energies above 10^{20} eV [17] and can be described by a sequence of power law spectra ($E^{-\alpha}$). The spectrum shows two distinct features caused by different spectral indices α. Below the *knee* at ≈ 2.5 PeV, the spectrum is described by a power law index $\alpha \approx 2.7$. Current theories point to cosmic rays from supernova remnants to be the dominant source [18, 19]. The *ankle* at ≈ 3 EeV marks the onset of extragalactic cosmic rays with $\alpha \approx 2.75$, possibly from giant radio galaxies, active galactic nuclei (AGN) and gamma ray bursts [20, 21]. The spectrum is steeper between the two features with $\alpha \approx 3.1$ [5], where also a weaker second knee is discussed. The upper limit of the spectrum is given by the Greisen–Zatsepin–Kuzmin

(GZK) limit[5], predicting a flux suppression at $\approx 10^{20}\,\mathrm{eV}$ by photo pion production on the cosmic microwave background [22, 23]. The cutoff has been confirmed by Auger and HiRes [24, 25]. Another explanation for this cutoff could be acceleration limits.

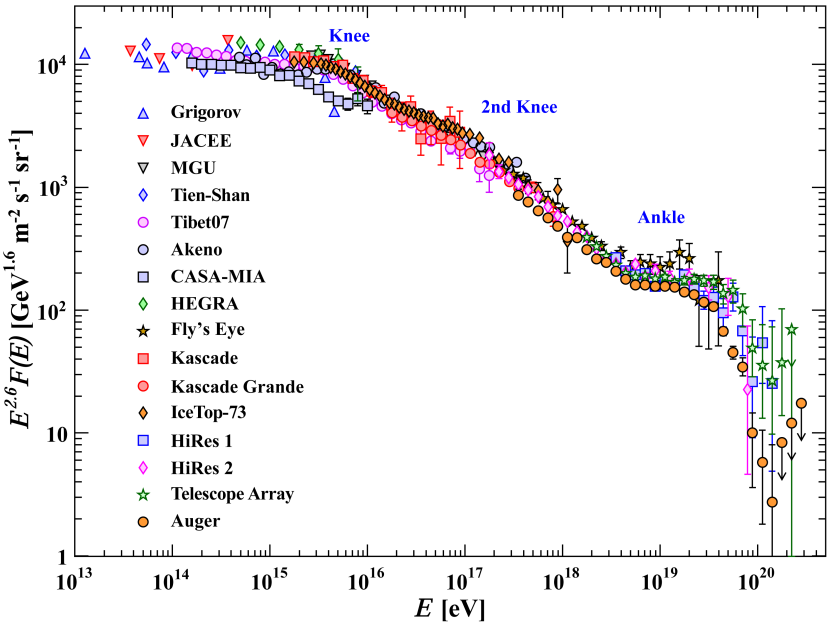


Figure 2.1: Cosmic ray spectrum from [16] as measured by the indicated experiments. The diffuse flux $F(E)$ has been multiplied by $E^{2.6}$ to make the spectral features more visible.

The mechanism how these cosmic rays are actually accelerated and how a universal power law spectrum emerges from potentially very different sources, was first covered by Enrico Fermi [26]. Particles would gain energy in head-on collisions with interstellar clouds and lose energy in following collisions. With the slightly higher probability for head-on collisions for moving particles, a net gain in energy is the result. The energy gain for a particle with speed v is proportional to $(v/c)^2$ and thus the mechanism was later named *second order Fermi acceleration*. Fermi showed that the resulting spectrum would be an inverse power law. However, the spectral index is not universal but depends on the characteristics of the acceleration site and the acceleration process would overall be very inefficient due to the small energy gain [27]. A solution was developed by Tony Bell applying the method to astrophysical shocks, e.g. from supernova remnants or active galactic nuclei jets [28]. A particle in its rest frame – being reflected back and forth through the shock front – experiences a compressive flow from both sides of the shock. This process results in head-on collisions only and an energy gain proportional to v/c (*first order Fermi acceleration*). If the relative energy gain per collision is given by ϵ and the probability to stay in the acceleration zone is P, then after k collisions, $N = N_0 P^k$ cosmic rays with an energy $E = E_0 \epsilon^k$ are produced. Combining the two expressions delivers [27]

[5]The cosmic ray suppression at highest energies was predicted by Kenneth Greisen, Vadim Kuzmin, and Georgiy Zatsepin at the same time.

$$\frac{\ln(N/N0)}{\ln(E/E0)} = \frac{\ln P}{\ln \epsilon}\,, \tag{2.9}$$

$$N = N_0 \left(\frac{E}{E_0}\right)^{\ln p/\ln \epsilon}\,. \tag{2.10}$$

Here N gives the number of particles with energies equal or larger than E and thus Equation 2.10 can be written in differential form as [27]

$$N(E)dE = \text{const.} \cdot E^{-1+(\ln p/\ln \epsilon)}\,. \tag{2.11}$$

From kinetic calculations one finds $\ln P/\ln \epsilon = -1$ [27]. Hence, first order Fermi acceleration provides an efficient acceleration mechanism, naturally resulting in a universal E^{-2} spectrum[6].

Although the shape of the spectrum is well known and a plausible acceleration method has been established, the exact origin of cosmic rays remains ambiguous. This is a consequence of the random walk of the charged particles due to deflection in the galactic magnetic field. Thus, when detected on earth, cosmic rays below the ankle carry no directional information about their production site. An exception is made by the highest energies around the GZK cutoff, where the deflection is only at the order of degrees [29]. Although the mean free path is reduced to $\approx 50\,\text{Mpc}$ by the GZK cutoff and statistics at these energies are very low, hints for a correlation of the directions with the closest active galactic nuclei were found by Auger [30].

Possible sources of particles with an energy E_{max} can be constrained by geometrical considerations, as the particle needs to be kept inside the acceleration region. Thus, the Larmor radius – depending on the strength of the magnetic field B – has to be smaller than the acceleration region R [21]. This limits the accessible energy to $E_{\text{max}} \leq \gamma eZBR$, with the particle charge eZ and the Lorentz factor γ. Figure 2.2 shows potential sources for cosmic rays at $10^{20}\,\text{eV}$. From this argument it becomes clear that the highest energy cosmic rays must be of extragalactic origin.

In the end, unequivocal evidence for the origin of cosmic rays at all energies will have to be revealed by alternative messengers. When protons are accelerated to sufficient energies, they can react with ambient baryons and photons right at the acceleration site producing neutral and charged pions (see Section 2.3.1.1 for interaction details). Gamma rays and neutrinos from pion decays are not deflected by magnetic fields and directly point towards their origin. However, photons still can be absorbed by dust and are strongly attenuated by pair production on the extragalactic background light above 0.1 TeV, as later shown in Figure 5.1. Additionally, gamma radiation can also be produced by other processes not connected to cosmic ray production, like inverse Compton scattering of relativistic electrons, superposing the neutral pion decay. Nevertheless, in rare cases, the neutral pion

[6]Note that the spectral index of $\alpha = 2$ results from an ideal calculation neglecting possible energy losses.

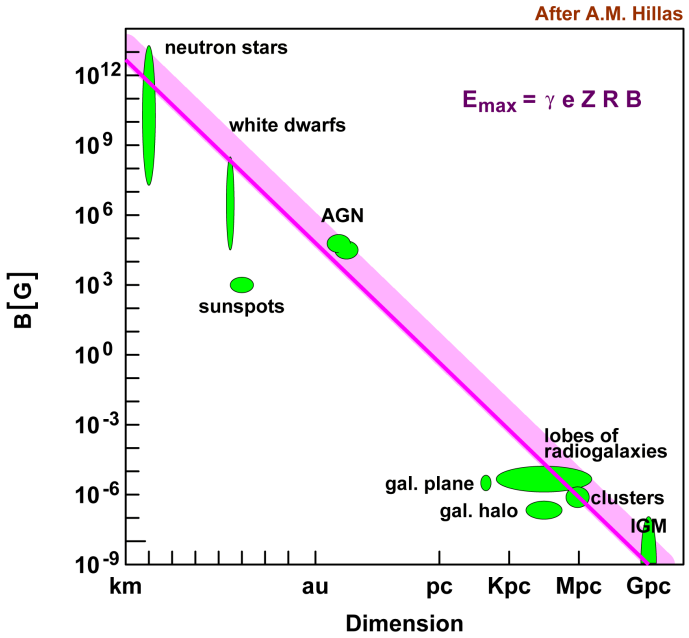


Figure 2.2: Hillas criterion: Objects above and within the shaded area in the spacial extent - magnetic field plane can constrain and accelerate particles up to 10^{20} eV. The lower edge of the shaded area represents the limit for iron nuclei and the upper edge the one for protons. IGM denotes the intergalactic medium. Modification from Hillas' original plot [21] from [20].

decay signature could be identified, for instance from the supernova remnants IC 443 and W44 by Fermi LAT [31].

Ultimately, the clearest evidence for cosmic ray acceleration sites would be the detection of a neutrino flux, as high-energy neutrinos are only produced in interactions of cosmic rays. With their low cross section and neutral charge, neutrinos are neither absorbed, nor deflected and thus are ideal messengers. However, because of these properties, large detectors of cubic kilometer scale are required for neutrino detection. The basic principles of underground/underwater neutrino telescopes have been proposed by Moisey Markov [32] and are covered in Section 2.4 after possible neutrino sources are described in Section 2.3.

2.3 Possible sources of astrophysical neutrinos

High-energy neutrinos are produced in the vicinity of astrophysical shocks where accelerated cosmic rays interact with ambient photons or baryons. Possible galactic sources are summarized in [5] and are expected to reach neutrino energies up to $\approx 100\,\mathrm{TeV}$.

Supernovae: At the end of the life cycle of massive stars, the missing fusion pressure triggers a gravitational collapse. The resulting supernova explosions emit total energies around $10^{46}\,\mathrm{J}$, of which 16 % are carried by anti electron neutrinos with energies around

15 MeV [33]. Such neutrinos have been detected e.g. by Kamiokande-II [34] from core collapse supernova SN 1987A, marking the first observation of astrophysical neutrinos.

Regarding high-energy neutrinos and cosmic rays, protons might be accelerated to energies above 10 TeV in the shock front of Type II (core collapse) supernovae. A TeV-neutrino flash of one hour duration would be created from proton-nucleon interactions [35].

Supernova remnants: The ejected shells from supernova explosions expand into space and form the supernova remnant. In recent years, observations revealed a class of supernovae surrounded by massive circumstellar material with densities of $\approx 10^5 - 10^{11}\mathrm{cm}^{-3}$. When the ejecta collide with the surrounding material, both forward and reverse shocks form and accelerate cosmic rays. With the massive surrounding target material, a TeV-neutrino flux over several months can be predicted [36].

Without the massive circumstellar material, supernova remnants still dominate the cosmic ray spectrum below the knee, but the neutrino production efficiency might be too low to detect single sources in cubic kilometer scale neutrino telescopes [5].

Molecular clouds and diffuse TeV-**gamma emission regions:** Molecular clouds with densities around $250\,\mathrm{cm}^{-3}$ can serve as a neutrino production site if a cosmic ray accelerator is located behind the cloud [5]. This is indicated by the observation of accompanying TeV-gamma ray emissions by H.E.S.S. – that possibly arise from proton-proton and proton-photon interactions [37]. The same processes might explain the TeV-gamma radiation from the Cygnus region as observed by the Milagro experiment [38]. About ten neutrinos detected per year with cubic kilometer neutrino telescopes are predicted in [39] from these sources.

Soft gamma repeaters: Neutron stars with extreme magnetic fields in the order of 10^{12} G emit frequent sub-second bursts of gamma radiation in irregular intervals. The emission is potentially triggered by reformations of the neutron star crust and its magnetic field [40]. Protons could be accelerated in the vicinity of such colossal magnetic fields, resulting in an observable neutrino flux from the strongest gamma flares such as from SGR 1806-20 in 2004 [41]. If the emission is observed in the hard X-ray region, these objects are also called anomalous X-ray pulsars.

Microquasars: Microquasars can be seen as a small variant of active galactic nuclei with a solar mass neutron star or black hole in its center. In massive binary systems, an accretion disk is formed by distracting matter from an accompanying massive star (Wolf-Rayet, B- or O-type [5]). Accelerated nuclei can interact in the accretion disk, with the compact object or with the accompanying massive star. Together with the TeV-gamma emission observed from these objects [5], this indicates a neutrino production and hence the possibility to detect several neutrinos per year [42].

As discussed before, the highest cosmic ray energies can only emerge from extragalactic sources. The two main extragalactic object classes in the context of cosmic ray acceleration and high-energy neutrino production are active galactic nuclei and gamma ray bursts which will be discussed below. As this work involves an analysis of blazars, the parent population of active galactic nuclei will be covered in more detail in Section 2.3.1.

Gamma ray bursts: In the 1960s, Vela satellites – launched to monitor possible nuclear weapon tests – detected several outbreaks of gamma radiation which did not originate from the earth or the sun, and neither showed the signature of weapon tests [43]. In subsequent observations, two subclasses have been identified by duration and spectral shape [44]. Short bursts with a duration below 2 s tend to show harder spectra and could be triggered by two merging compact objects (two neutron stars or a neutron star and a black hole) [45]. Longer bursts, typically around 2 s emit softer gamma spectra and could originate from *failed* supernova collapsing into a black hole [46]. An afterglow from lower energies down to radio emission can sometimes still be observed years after the burst [47].

Given the large distances with redshifts up to $z = 8$ [48] in which gamma ray bursts are observed, they are the intrinsically brightest objects in the universe with a total electromagnetic radiation output of $\approx 10^{44}$ J (about 200×earth mass) [49]. The development, cosmic ray acceleration and neutrino production from these events is commonly described in the so called fireball model [50]. Recent IceCube results strongly constrain the parameter space for these models, though [51].

2.3.1 Active galactic nuclei

Active galactic nuclei (AGNs) are distant extragalactic point-like sources of bright electromagnetic radiation. They have broad emission characteristics ranging from radio to TeV gamma rays and have been identified to reside in the small[7] centers of their host galaxies (active galaxies). Although AGNs can outshine the entire surrounding disk of billions of stars by orders of magnitude, the host galaxy can be identified in many cases, revealing spiral, irregular and gigantic elliptical shapes [53, 54]. Being the most powerful persistent objects in the universe with intrinsic γ-luminosities reaching 10^{42} W [55] and observed redshifts up to $z = 7$ [56], AGNs are not only candidates for the origin of high-energy cosmic rays but also shed light on the evolution of the universe on large scales[8].

Historically, many different AGN classes were identified based on observations in different frequency bands creating a somewhat confusing zoo of objects and names. Below, these AGN types, also listed in Table 2.1, are described systematically before we introduce a unified AGN model. First, AGNs can be divided into radio-loud[9] objects (mostly in the centers of elliptical host galaxies with typically large scale radio jets) and radio-quiet objects (mostly situated in spiral host galaxies) [54].

[7] The rapid variability of blazars limits the size of the emission region to < 1 light day [52].

[8] For example, the most distant detected blazar ULAS J112001.48+064124 at $z = 7.085$ (0.77 billion years after the Big Bang) allows the examination of the reionization period of the universe [56].

[9] A distinction is achieved by the radio (5 GHz) to optical (4400 Å) ratio [57].

Emission lines	Radio loud		Radio quiet
	FR I	FR II	
broad	-	FSRQ, Quasar, SSRQ, BLRG	Seyfert 1 QSO
narrow or none	BL Lac, NLRG, Radio galaxy	NLRG, Radio galaxy	Seyfert 2

Table 2.1: AGN types and emission properties. Abbreviations are explained in the text.

The radio-quiet class was first examined by Carl Seyfert. In optical and ultraviolet (UV) spectroscopies of spiral active galaxies he found objects showing broad[10] and narrow emission lines (Seyfert 1) as well as objects lacking the broad lines (Seyfert 2) [53]. High luminosity Seyfert 1 galaxies are named Quasi Stellar Objects (QSO).

The sequence of Quasars, Steep Spectrum Radio Quasars (SSRQ) and Broad-Line Radio Galaxies (BLRG) from high to low luminosity can be seen as radio-loud counterpart of Seyfert 1 galaxies because of their broad emission lines. The radio-loud objects corresponding to the Seyfert 2 class are called Narrow-Line Radio Galaxies (NLRG) or just Radio Galaxies.

Later, Fanaroff and Riley analyzed the relativistic radio jets within the radio-loud group and found two separated populations. Luminous radio-loud AGNs show a jet emission maximum further away from the galaxy (FR II) while the jet brightness declines with the distance in less luminous AGNs (FR I) [59]. Additionally, FR II AGNs show an order of magnitude stronger optical emission lines compared to FR I AGNs [54].

Within the radio-loud population, objects with rapid variability, extremely high γ-luminosities, compact radio cores, and strong radiation polarization are called blazars. They appear with broad emission lines as a subclass of FR II galaxies named Flat Spectrum Radio Quasars (FSRQ), as well as with a weak or missing broad line region as BL Lac (FR I) [54] (Fig. 2.4). A commonly used criterion to divide the blazar subclasses requires an equivalent emission line width in the rest frame of 5 Å for FSRQs [61].

Today, this wide range of emission characteristics is explained by a highly anisotropic emission profile and, consequently, a strong emission dependence on the observation angle. The widely accepted unified model (Fig. 2.3) assumes a supermassive black hole surrounded by an accretion disc of infalling matter in the center of a host galaxy. The incoming material is heating up and glowing brightly in the optical/UV band [62]. Hot, excited gas surrounding the central engine forms the broad line region (black blobs in Fig. 2.3), while narrow emission lines originate from colder, more distant clouds (gray blobs in Fig. 2.3). A torus of dust prevents the transverse observation of the central engine and the broad line region. Perpendicular to the disk, relativistic radio jets emanate from the black hole reaching extents in the order of 10^6 light years [54]. These jets are possibly powered and

[10]The width of emission lines is directly connected to the speed/temperature of the emitting atoms by the Doppler effect.

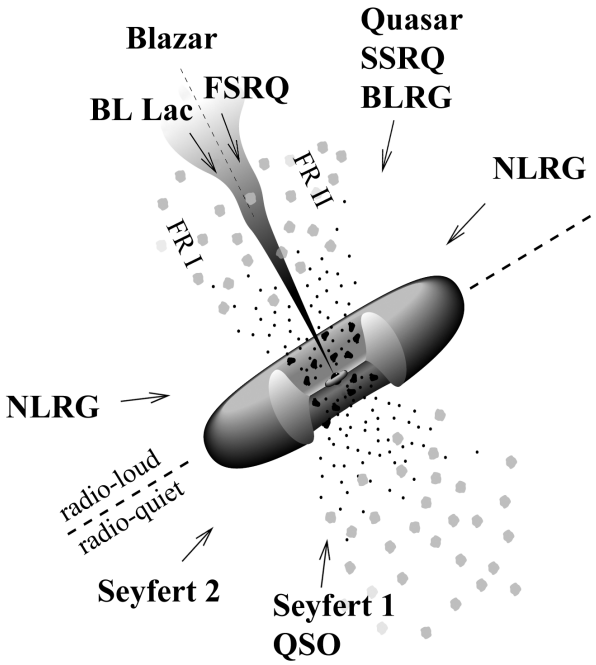


Figure 2.3: In the unified model, AGNs are powered by matter accretion in a disk surrounding a supermassive black hole. Relativistic radio jets are emitted perpendicular to the disk in the radio-loud case. Broad emission lines originate from hot gas close to the center (broad line region, black blobs) while narrow lines are emitted from colder distant clouds (gray blobs). On the one hand, the observed emission profile heavily depends on the intrinsic object type (radio-loud/quiet, FR I/II). On the other hand, it depends on the observation angle and whether the inner engine and broad line region are obscured by the torus of dust (NLRG, Radio Galaxy), or can be directly observed (Quasar, SSRQ, BLRG, Seyfert 1, QSO), or are even viewed along the radio jet (Blazars: FSRQ, BL Lac). Image modified from [58].

beamed by electromagnetic extraction of energy and angular momentum from the black hole via the Blandford-Znajek process [63].

In this model, the lack of broad emission lines from NLRGs and Seyfert 2 galaxies is due to absorption in the dust torus around the central engine, while FR I AGNs are assumed to have an intrinsically weak broad line region. The apparent radio luminosity could be, among other characteristics, connected to the black hole mass, spin, accretion rate, and host galaxy type [54].

The emission characteristics of blazars are explained by the line of sight being aligned with the radio jet [54, 55]. The extreme γ-luminosities in the GeV and TeV band suggest that blazars could be some of the strongest accelerators of relativistic particles in the universe. Fermi acceleration due to shocks in the relativistic radio jets might produce ultra high-energy cosmic rays (UHECR) [29, 55, 64, 65, 66, 67]. The detection of high-energy neutrinos from blazars would support this scenario and prove hadronic acceleration. In the alternative AGN emission from pure leptonic processes, no neutrinos would be produced. These two classes of blazar emission models and the connection to UHECR and neutrinos are discussed in Section 2.3.1.1.

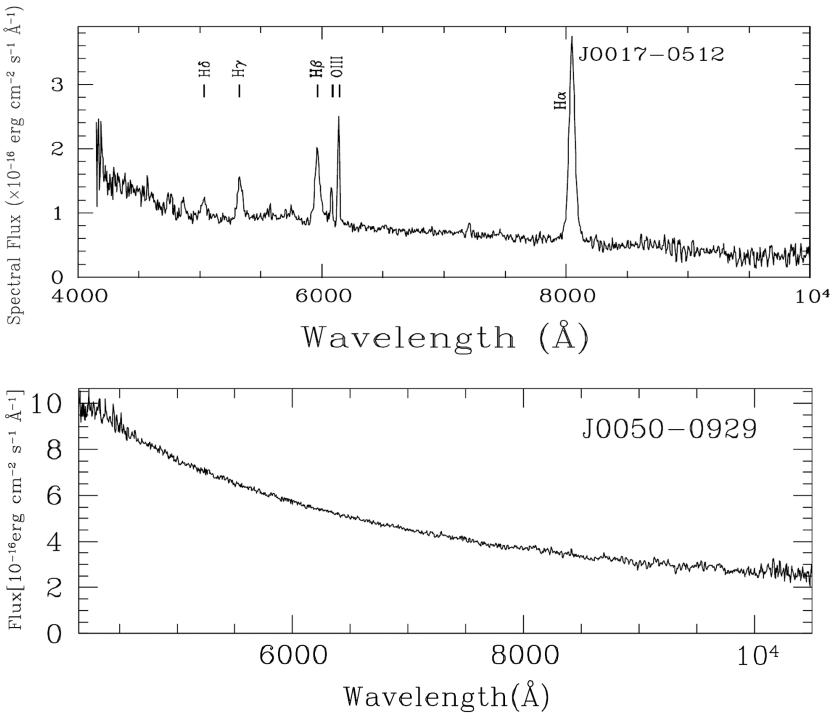


Figure 2.4: The upper plot shows a typical optical spectrum of FSRQs with broad emission lines which are absent in BL Lac spectra as demonstrated in the lower plot [60].

2.3.1.1 Blazar emission models

The spectral energy distribution of blazars basically consists of a low energy bump from radio to X-ray energies and a high-energy component from X-ray up to γ-ray energies (Fig. 2.5). The peak frequency f_{peak} of the lower part of the spectrum, commonly explained by synchrotron radiation of relativistic electrons [52, 68], further classifies blazars as being low- ($f_{peak} < 10^{14}$ Hz), intermediate- (10^{14} Hz $< f_{peak} < 10^{15}$ Hz) or high- ($f_{peak} > 10^{15}$ Hz) synchrotron peaked (LSP, ISP, HSP) [69].

The γ-emission could, in contrast, originate from pure leptonic acceleration or lepto-hadronic processes [52, 68]. In both cases a realistic blazar model also needs to reproduce the observed variability and time dependence within the spectral energy distribution.

Leptonic models: In leptonic models, the high-energy feature in the spectral energy distribution of blazars is generated by the inverse Compton effect. Soft photons are up-scattered to gamma-ray energies by relativistic electrons accelerated in the jet. The soft photon field can be the synchrotron radiation from the very electrons within the jet (self synchrotron Compton, SSC) or it can originate from external sources (external Compton, EC). Possible sources outside the jet are the accretion disk, the broad line region, or infrared emission from the dust torus [71]. While SSC models successfully replicate the electromagnetic observations from BL Lacs, the more luminous FSRQs sometimes require external photons, more sophisticated jet topologies and/or hadronic acceleration [71, 72].

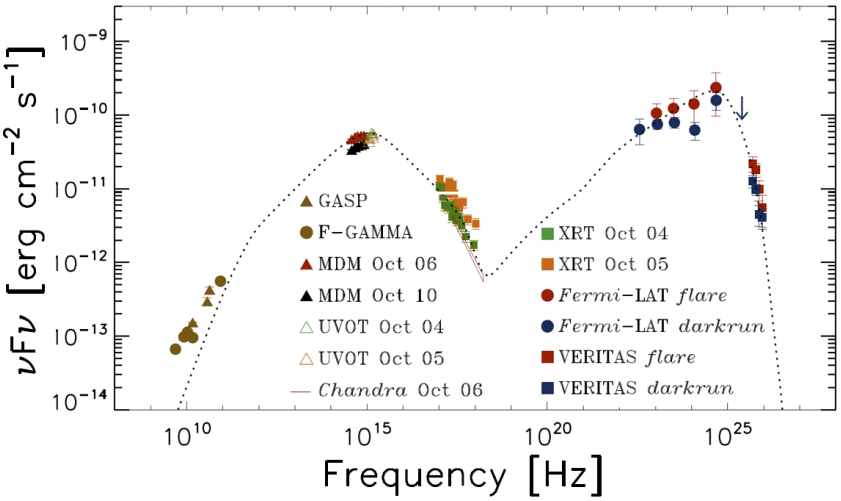


Figure 2.5: Spectral energy distribution of the BL Lac 3C 66A from a multiwavelength observation in October 2008 [70]. Here, $\nu F\nu$ stands for the observed differential flux at earth weighted by the square of the energy. A leptonic blazar model combining EC emission and SSC emission is show as a dashed line.

Hadronic models: If protons are accelerated above the photo-pion interaction threshold in the jet, a significant fraction of the γ-emission could originate from π^0-decays [64, 68]. Equation 2.12 shows the interaction chain of the dominant Δ-excitation, having a branching ratio of 2/3 to 1/3 in favor of π^0 and a Δ cross section of $\sigma_\Delta \approx 10^{-28}\,\mathrm{cm}^2$ [73, 74]:

$$p\gamma \longrightarrow \Delta^+ \longrightarrow \begin{cases} n\pi^+ & \longrightarrow n\mu^+\nu_\mu \longrightarrow ne^+\nu_e\nu_\mu\bar{\nu}_\mu \\ p\pi^0 & \longrightarrow p\gamma\gamma \end{cases} . \tag{2.12}$$

A second variant of hadronic models considers proton-proton interactions within the jet, also leading to pion production. High-energy protons can interact with low energy protons according to [5, 75]

$$pp \longrightarrow \begin{cases} pn\pi^+ & \longrightarrow pn\mu^+\nu_\mu \longrightarrow pne^+\nu_e\nu_\mu\bar{\nu}_\mu \\ pp\pi^0 & \longrightarrow pp\gamma\gamma \end{cases} , \tag{2.13}$$

again with the same branching ratios 2/3 to 1/3 favoring the π^0 production [5]. As stated in Section 2.1, these processes generate a flavor ratio of $(\nu_e : \nu_\mu : \nu_\tau) = (1 : 2 : 0)$ which results in an oscillated ratio of $\approx (1 : 1 : 1)$ at earth.

Hadronic models generally require strong magnetic fields B of several tens of Gauss to keep the proton Larmor radius $R_L \propto E_p B^{-1}$ smaller than the size of the emission region, such that protons are accelerated above the required energies [52, 68]. In the vicinity of such fields, synchrotron radiation from secondary muons, pions and electrons, and from the primary protons has to be considered for a self-consistent synchrotron proton blazar

model (SPB) [52, 68, 75]. SPB models have been demonstrated to successfully reproduce the spectral energy distribution of different BL Lac and FSRQ types [64, 68, 75]. The detection of TeV neutrinos from the decay of charged pions, produced along with neutral pions, would confirm this scenario.

2.4 High-energy neutrino detection

Neutrinos can be detected on the basis of their rare interactions with matter. The main channels for high-energy neutrino detection are charged current (CC) and neutral current (NC) interactions with an atomic nucleus N [5]:

$$\text{NC:} \quad \nu_l + N \longrightarrow \nu_l + H\,, \tag{2.14}$$

$$\text{CC:} \quad \nu_l + N \longrightarrow l + H\,, \tag{2.15}$$

with the lepton/neutrino flavor l (e, μ, τ) and a hadronic cascade H. The resulting charged particles emit Cherenkov radiation (Section 2.4.2), which can be detected by photomultiplier tubes (PMT) in neutrino telescopes (Section 2.5). The light signature depends on the interaction type and the neutrino flavor. Hadronic cascades reach a maximum extent of $\approx 10\,\text{m}$ [76] and thus cannot be resolved in the coarse grids of high-energy neutrino telescopes. Thus, all NC interactions appear as point-like, almost isotropic light sources (cascade in Figure 2.6). The typical range of electromagnetic cascades, created in CC interactions of electron neutrinos, is of the same order and thus produces a similar light signature.

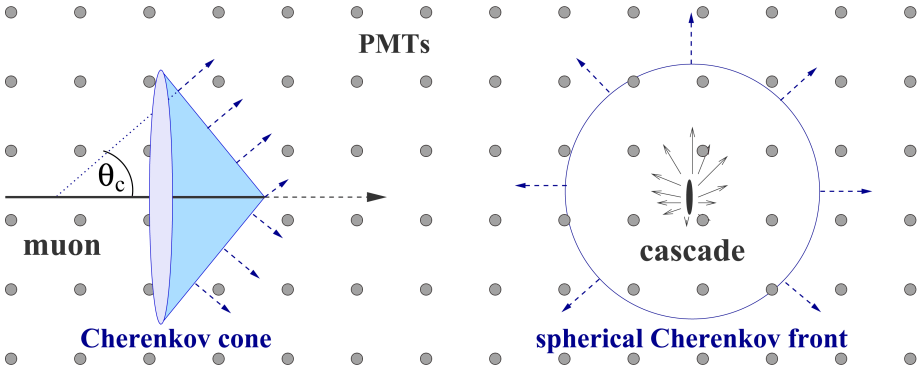


Figure 2.6: Cherenkov light signatures for muons on the left and cascades on the right. Plot from [77].

The light pattern of a ν_τ CC event strongly depends on the primary neutrino energy and the range of the subsequent tau lepton. With the extremely short τ lifetime of $2.9 \times 10^{-13}\,\text{s}$ [78], the average range rises linearly with energy [76] ($\approx 50\,\text{m} \cdot E_{\nu_\tau}[\text{PeV}]$). Below $\approx 1\,\text{PeV}$, the immediate decay results in an almost point-like light source that can

only be distinguished by identification of double pulse signatures in the individual PMT waveforms, while above that energy the energy deposition from the CC interaction vertex and the tau lepton decay vertex may be resolved as so called *double bang* signature. All cascade-like light signatures described here only allow a poor angular reconstruction of 10° or worse in cubic kilometer neutrino telescopes because of their close to isotropic light pattern.

However, muons from CC muon neutrino interactions already create long tracks at GeV-energies as shown in Figures 2.6 and 2.9. The dimensions of cubic kilometer scale detectors are exceeded at 1 TeV therefore extending the effective volume, as muons from neutrino interactions outside the detector still can enter the instrumented volume. The muon propagation is covered in greater detail in Section 2.4.1. Because of the long muon tracks, an angular reconstruction with accuracies down to 0.4° at highest energies is possible in the CC ν_μ-channel as demonstrated in Chapter 3. The interaction angle ψ between the muon neutrino and the muon depends on the neutrino energy E_ν and is given by [79]

$$\psi = 0.7^\circ \cdot (E_\nu/\mathrm{TeV})^{-0.7}\,, \tag{2.16}$$

which is below the achieved muon reconstruction accuracy at all energies. Hence, the CC ν_μ-channel allows for the search for point-sources of astrophysical neutrinos.

The detection probability for high-energy neutrinos depends on the neutrino nucleon cross section. Only the charged current cross section from [80] is given here, as it is the relevant interaction channel for this work. The differential cross section can be written as

$$\frac{d^2\sigma}{dxdy} = \frac{2G_F^2 M E_\nu}{\pi}\left(\frac{M_W^2}{Q^2+M_W^2}\right)^2\left[xq(x,Q^2)+x\bar{q}(x,Q^2)(1-y)^2\right], \tag{2.17}$$

with the Bjorken scaling variables $x = Q^2/2M\nu$ and $y = \nu/E_\nu$, the invariant momentum transfer $-Q^2$ between neutrino and muon, the energy loss $\nu = E_\nu - E_\mu$ in the target frame, the nucleon and W-Boson masses M and M_W, and the Fermi constant $G_F = 1.166\,32 \times 10^{-5}\,\mathrm{GeV}^{-2}$. The quark distributions are given by

$$\begin{aligned} q\left(x,Q^2\right) &= \frac{u_v\left(x,Q^2\right)+d_v\left(x,Q^2\right)}{2} + \frac{u_s\left(x,Q^2\right)+d_s\left(x,Q^2\right)}{2} \\ &\quad + s_s\left(x,Q^2\right) + b_s\left(x,Q^2\right)\,, \end{aligned} \tag{2.18}$$

$$\bar{q}\left(x,Q^2\right) = \frac{u_s\left(x,Q^2\right)+d_s\left(x,Q^2\right)}{2} + c_s\left(x,Q^2\right) + t_s\left(x,Q^2\right)\,, \tag{2.19}$$

with sea and valence quarks denoted by s and v and the quark flavors u, d, c, s, t, b. NC and CC cross sections are shown on the left of Figure 2.7 as function of the neutrino energy. Below 10^4 GeV, the CC cross section rises linearly, while it is damped by the W-boson above that energy. Above 10^6 GeV, ν and $\bar{\nu}$ show equal cross sections, as the valence quark contribution becomes negligible [80]. A more demonstrative representation can be given by the interaction length [80]

$$L_{\text{int}} = \frac{1}{\sigma_{\nu N}(E_\nu) N_A}\,, \tag{2.20}$$

with the Avogadro number $N_A = 6.022 \times 10^{23}\,\text{mol}^{-1} = 6.022 \times 10^{23}\,\text{cm}^{-3}$ (water equivalent). The resulting interaction lengths for water are shown in the right plot in Figure 2.7. It is apparent that the mean free path drops below the diameter of the earth around 100 TeV. Hence, the earth becomes opaque for neutrinos above that energy and the neutrino detection at the highest energies is restricted to the horizon and downgoing tracks.

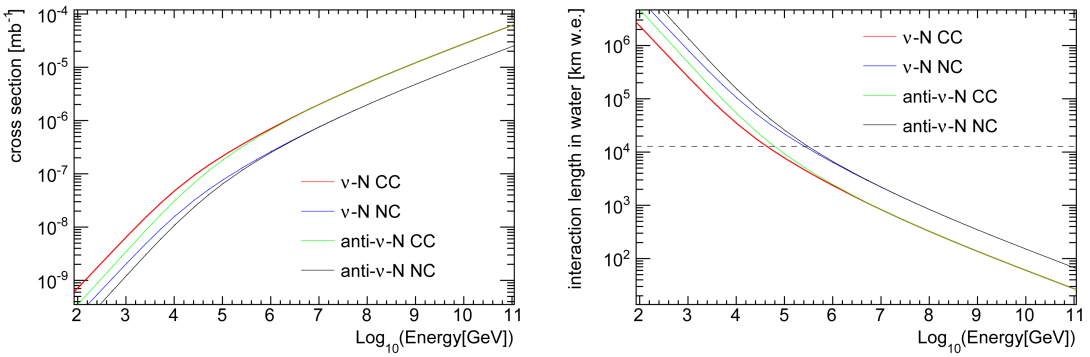


Figure 2.7: Neutrino cross sections and interaction lengths (water equivalent w.e.) both depending on the neutrino energy. The diameter of the earth is shown as dashed line. Plot from [49].

With the given cross sections, it is apparent that huge detection volumes are necessary to obtain sufficient statistics. With the required optical transparency for Cherenkov light detection, only natural ice or water reservoirs are viable detector sites. This section will focus on the neutrino detection in ice with IceCube.

2.4.1 Muon propagation

As stated in Section 2.4, neutrino point source searches are based on muons from muon neutrino interactions. Key to the propagation of muons through matter is the understanding of the different energy loss processes. The continuous emission of Cherenkov radiation (Section 2.4.2) only accounts for a tiny fraction of the total energy loss and can be ignored in this context [81]. A muon mainly looses energy by ionization, bremsstrahlung, e^+e^--pair production and photo-nuclear interactions. The energy loss per meter dE/dx from ionization is nearly constant while the other radiation losses approximately rise linearly with the muon energy E_μ [82]. Thus, the energy loss can be expressed as [83]

$$-\frac{dE}{dx} = A + BE_\mu\,, \tag{2.21}$$

with $A = 0.24\,\mathrm{GeV\,m^{-1}}$ and $B = 0.000\,32\,\mathrm{m^{-1}}$ for ice [81]. The radiative losses equal the ionization losses at a critical energy $A/B = 750\,\mathrm{GeV}$. In Figure 2.8, displaying dE/dx vs. E_μ, the dominance of radiative losses above the critical energy becomes obvious. Hence, because of the stochastic nature of the radiative losses, only the average muon energy $\langle E_\mu \rangle$ after x meters and the average range $\langle R(E_0) \rangle$ of a muon can be evaluated by solving the differential Equation 2.21, giving

$$\langle E_\mu(x) \rangle = \frac{e^{-Bx} \cdot (A - A \cdot e^{Bx} + BE_0)}{B}\,, \tag{2.22}$$

$$\langle R(E_0) \rangle = \frac{\log\left(\frac{A+BE_0}{A}\right)}{B}\,, \tag{2.23}$$

with the initial muon energy E_0. Figure 2.9 shows a typical range of $\approx 2\,\mathrm{km}$ for a muon with $E_0 = 1\,\mathrm{TeV}$.

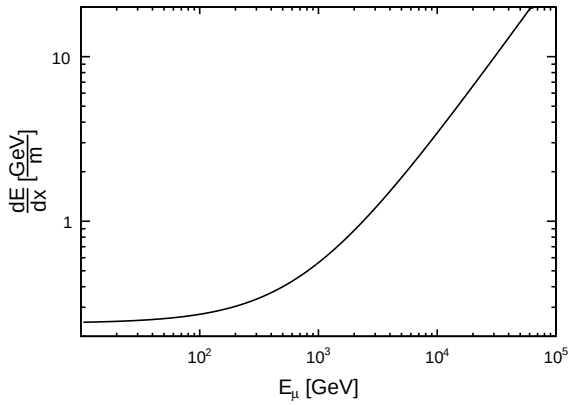


Figure 2.8: Average muon energy loss per meter dE/dx vs. muon energy E_μ

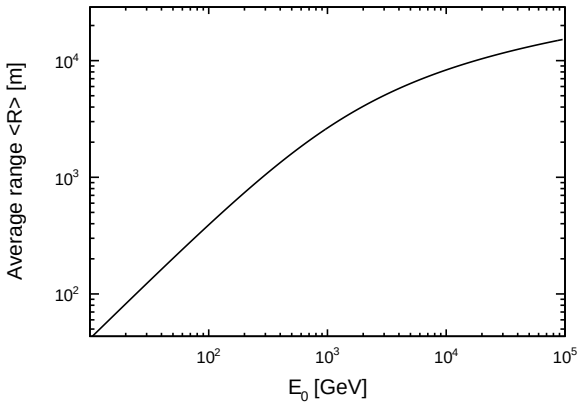


Figure 2.9: Average muon range $\langle R \rangle$ in ice depending on the initial muon energy E_0.

The contributions from the particular energy loss processes are e.g. given in [82] and are the foundation for the simulation of the muon propagation covered in Section 2.5.5.3.

2.4.2 The Cherenkov effect

When charged particles travel through matter faster than the phase velocity of light in the medium, they emit Cherenkov radiation[11]. Thus the Cherenkov threshold is defined by the phase velocity spectral index n_p as

$$\beta \geq \frac{1}{n_p}\,, \tag{2.24}$$

[11] First observed by Pavel Alekseyevich Cherenkov in 1937 [84].

with $\beta^2 = v^2/c^2$. The spectrum of the radiation is given as number of photons N per unit length x and wavelength λ by the Frank-Tamm formula [85]

$$\frac{d^2N}{dxd\lambda} = \frac{2\pi\alpha}{\lambda^2} \cdot \left(1 - \frac{1}{n^2\beta^2}\right) , \tag{2.25}$$

with the fine-structure constant $\alpha \approx 1/137$. The energy loss due to Cherenkov radiation is about $0.1\,\mathrm{MeV\,m^{-1}}$ [86] and thus negligible compared to the contributions discussed in Section 2.4.1. The photon emission angle depends on the phase velocity refraction index and is given by [86]

$$\cos\left(\Theta_C\right) = \frac{1}{\beta n_p} . \tag{2.26}$$

One finds $\Theta_C = 40.75°$ for relativistic particles with $\beta \approx 1$ and a refraction index for ice $n_p = 1.32$. With the fixed photon emission angle along high-energy muons, the muon track can be reconstructed from the detected light in transparent media. The light detection of an in-ice Cherenkov telescope has to be optimized to yield a good efficiency respecting the λ^{-2} Cherenkov spectrum and the optical ice properties.

2.4.3 Photon propagation in the deep Antarctic ice

Light propagation through ice or water and accurate knowledge of the optical properties of the medium is a key element for the understanding and performance of Cherenkov telescopes. This topic is covered in much detail in the Photonics propagation framework publication [87] and IceCube ice model proceedings [88].

The photon propagation in ice is characterized by the speed of light in the medium[12], scattering and absorption. Photons can be absorbed in excitation processes of electrons or molecules while deflection occurs on dust particles, ice crystal defects or air bubbles [87].

The scattering process on spherical particles is examined theoretically in the Mie theory [89] and depends on the mean free path or scattering length λ_s and on the scattering angle θ. The average cosine $g = \langle\cos\theta\rangle$ describes the nature of the process being isotropic for $\langle\cos\theta\rangle = 0$. For the Antarctic ice, a highly forward orientated distribution with g close to 1 is calculated from Mie theory [88]. The evaluated probability density for the scattering angle $f(\cos\theta)$ can be approximated by a linear combination of the Henyey-Greenstein function[13] [90] $\mathrm{HG}(\cos\theta)$ and a Liu scattering function [91] $\mathrm{SL}(\cos\theta)$ to [92]

[12] The speed of light in ice c_{ice} is calculated from the light speed in vacuum c_{vac} and the group velocity refraction index $n_{\mathrm{group}} = 1.35634$ as $c_{\mathrm{ice}} = c_{\mathrm{vac}}/n_{\mathrm{group}}$.

[13] The Henyey-Greenstein function was originally used to explain the diffuse interstellar radiation by starlight scattered on interstellar matter.

$$f(\cos\theta) = (1 - f_{\mathrm{SL}}) \cdot \mathrm{HG}(\cos\theta) + f_{\mathrm{SL}} \cdot \mathrm{SL}(\cos\theta) \quad \text{, with} \tag{2.27}$$

$$\mathrm{HG}(\cos\theta) = \frac{1 - g^2}{2\left(1 + g^2 - 2g\cos\theta\right)^{3/2}} \quad \text{and} \tag{2.28}$$

$$\mathrm{SL}(\cos\theta) = (1 + \cos\theta)^{\frac{2g}{1-g}} \quad . \tag{2.29}$$

Absorption is characterized by the absorption length λ_a, expressing the traveled distance after which the photon density decreases to $1/e$. Hence, the photon propagation relies on an ice model which has to provide the scattering and absorption coefficients as well as other free parameters like f_{SL} and $g = \langle\cos\theta\rangle$ at every photon position. The ice model used in IceCube is covered in Section 2.5.4.

2.4.4 Cosmic ray backgrounds

It can be seen as an ironic twist that, when looking for the origins of cosmic rays with neutrino telescopes, the main background is caused by the very cosmic rays. Hadronic interactions of cosmic rays with particles in the atmosphere constantly create large particle showers. The impact of a single cosmic ray particle can trigger cascades containing up to several thousand mesons depending on its energy [93]. These mesons are mainly kaons and pions produced by [94]

$$p + N \longrightarrow \pi^+ + \pi^- + K^+ + K^- + \dots \,, \tag{2.30}$$

for a cosmic ray proton p and an atmospheric nucleus N. Muons and neutrinos are produced in subsequent decays by [94]

$$\pi^+ \left(K^+\right) \longrightarrow \mu^+ + \nu_\mu \tag{2.31}$$

$$\pi^- \left(K^-\right) \longrightarrow \mu^- + \bar{\nu}_\mu \,. \tag{2.32}$$

Pions and kaons have rest frame lifetimes of $\approx 10^{-8}\,\mathrm{s}$ and thus can loose a fraction of their initial energy before decaying. As a result, the created atmospheric muon and neutrino spectra are steeper than the cosmic ray spectrum and develop a power law spectrum with an index $\alpha \approx 3.7$ [93, 94, 95]. Above 100 GeV, muons are mainly created in pion decays and muon neutrinos primarily originate from kaon decays [96]. Electron neutrinos from muon decays can be neglected in this context, as muons already reach the ground and the detector above energies of a few GeV [97]. The flux of the atmospheric particles additionally depends on the zenith angle due to different density gradients in the atmosphere [94]. Additional to these *conventional* atmospheric fluxes from kaons and pions, heavy hadrons with charm contribution (D^0, D^+, D_s^+, Λ_c^+) can be created at sufficient high energies [98]. With lifetimes below $10^{-12}\,\mathrm{s}$ these hadrons decay instantly and the resulting leptons are thus named *prompt* [93]. Because of the short lifetime, the hadrons do not suffer from

a significant energy loss and preserve the shape of the primary cosmic ray spectrum. Although there are large uncertainties in the cross sections for charm production, the prompt flux is expected to dominate the atmospheric muon neutrino spectrum at $E_\nu > 100\,\mathrm{TeV}$ [96]. Figures 2.10 and 2.11 show the atmospheric muon and neutrino spectrum. In both cases, the steep conventional flux and the harder prompt predictions at higher energies are clearly visible.

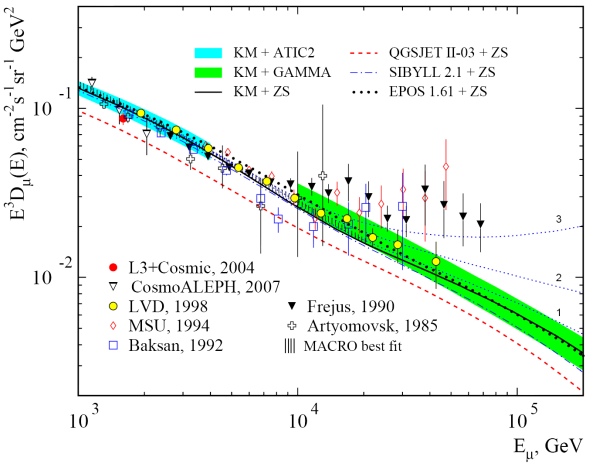


Figure 2.10: Vertical atmospheric muon flux at ground level for different conventional models (upper labels), prompt models (indicated by the numbers 1,2,3) and experimental data (lower labels). The models shown are further specified in [99]. Plot modification taken from [95].

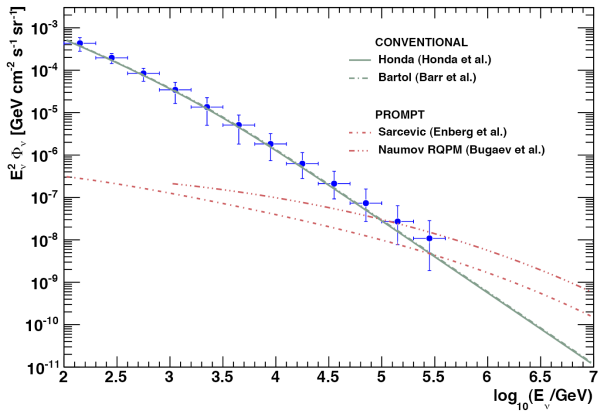


Figure 2.11: Average northern hemisphere atmospheric muon neutrino flux measured with IC40 [93] with conventional flux predictions by Honda [100] and Bartol [101] and prompt flux predictions by Enberg [98] and Naumov [102]. Plot taken from [94].

According to Equation 2.23, muons can maximally travel several tenths of km through ice and have an even shorter absorption length in the denser bedrock. This has consequences for arrival directions of different signal types in a Cherenkov telescope. In the case of IceCube – which is located at the geographic South Pole and described in Section 2.5 – atmospheric muons can enter the detector only from the southern hemisphere while muons induced by atmospheric neutrinos arrive from isotropic directions. This demands a different handling of the two background types and hemispheres in background rejection. The situation is sketched in Figure 2.12.

For the northern hemisphere, the main task is the identification of falsely reconstructed atmospheric muons mainly caused by two or more coincident tracks[14] in the detector mimicking an upgoing direction. These can be identified by reconstruction quality parameters or the investigation of the photon hit pattern. A potential astrophysical neutrino flux can then be identified in the remaining atmospheric neutrinos by a harder energy spectrum and/or a local excess. The searches for astrophysical neutrinos from the southern

[14]Coincidences can occur from different atmospheric particle showers or also involve both atmospheric or astrophysical neutrinos.

hemisphere are restricted to higher energies, where the strong but rather soft atmospheric muon flux fades away.

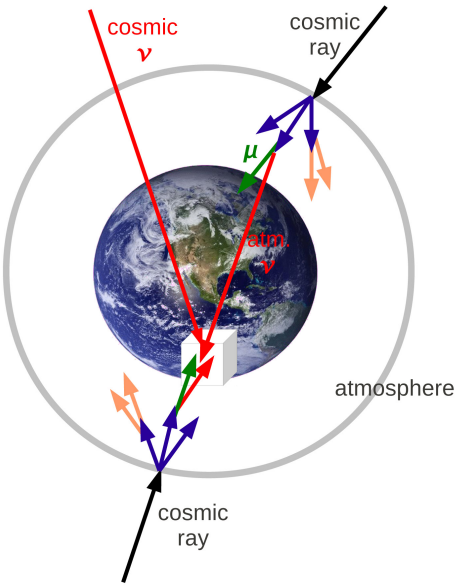


Figure 2.12: Different signal types in IceCube [49]. Neutrinos are shown red and can reach the detector from all directions. In contrast, muons, shown in green, are absorbed in the earth.

2.5 The IceCube neutrino telescope

IceCube is a high-energy neutrino telescope deployed into the 3 km thick ice sheet close to the geographic South Pole. It has been constructed in the Antarctic summer seasons between the years 2005 and 2010. Figure 2.13 shows the hexagonal detector design with its 78 holes (60 cm diameter) in a 125 m spacing created with a hot water drill[15]. Each hole contains a string with 60 digital optical modules (DOM) attached equidistantly between depths of 1450 m and 2450 m resulting in a vertical DOM to DOM distance of 17 m. The DOMs are pairwise connected to the surface with 30 twisted pair cables per string [103]. The optical modules are equipped with large area photomultiplier tubes (PMT) allowing the detection of Cherenkov radiation emitted by charged leptons traveling through the detector.

The low energy extension DeepCore consists of eight additional strings, six of them with high efficiency PMTs, and the bottom third of the seven central IceCube strings. The average horizontal string distance is 72 m and the vertical DOM spacing of the DeepCore-only strings amounts to 7 m. The denser instrumentation within the lower central part of

[15]The hot water drill with a power of 5 MW provides 200 gallons/minute of 88 °C hot water. Drilling takes about 40 h and string deployment about 12 h [103].

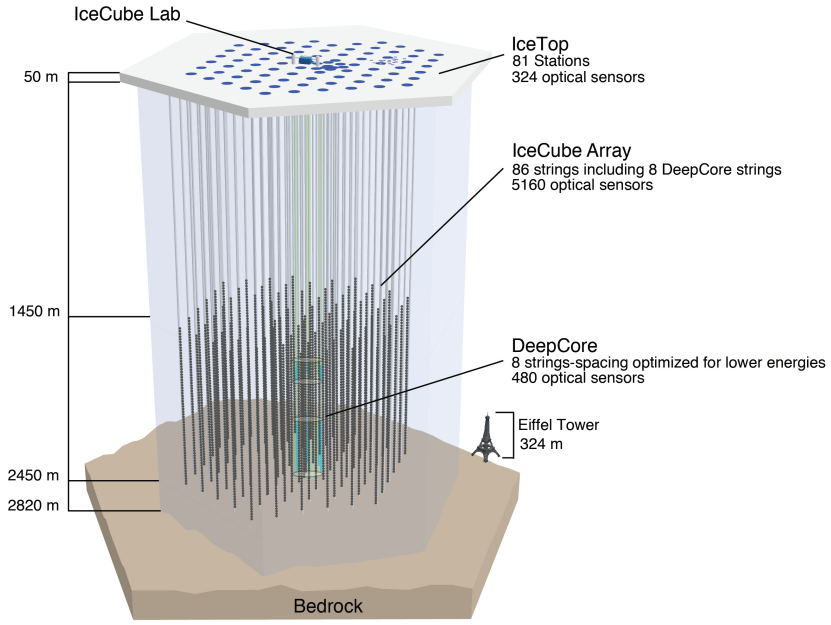


Figure 2.13: The IceCube detector: 86 strings, equipped with photomultiplier tubes, create an instrumented volume of one cubic kilometer ice in depths between 1450 m and 2450 m close to the geographic South Pole.

IceCube and the use of the surrounding detector as veto for atmospheric muons expands the energy range down by a factor of 10 towards $E_\nu \gtrsim 10\,\text{GeV}$ [104]. Finally, two ice tanks of 1.86 m diameter with two DOMs respectively at the surface of 81 holes form the IceTop array. It is used as a cosmic ray detector and veto for downgoing atmospheric muons within IceCube.

The data of all detector components is collected and processed at a surface station named IceCube Lab (ICL). The partial detector configurations before completion are named according to the number of deployed strings at that time as IC22, IC40, IC59, IC79 and finally IC86. As the detector was already operating and collecting data in these years, the resulting data sets are labeled accordingly.

2.5.1 The digital optical module (DOM)

IceCube's digital optical modules (Figure 2.14) are the fundamental components of the detector. Operating under extreme pressure and temperature conditions, they have to provide reliable high precision light measurements for years, not accessible for maintenance once being frozen. A DOM's PMT and electronics are enclosed by a 13 mm thick borosilicate glass sphere with a diameter of 33 cm. The housing has to resist the high pressure of up to 400 atm during refreeze [105].

The main DOM component is a Hamamatsu R7081-02 photomultiplier tube with a gain of 10^7, whose high voltage of 1500 V is provided by a Cockroft-Walton power supply [103].

The R7081-02 works with 10 dynode stages delivering an average signal of 8 mV per photon, defining one photo electron (pe) [106]. The photocathode has a diameter of 10 in, consists of standard bialkali materials and has a maximum quantum efficiency[16] of $\approx$ 25 % at 390 nm, which reasonably matches the maximum transparency of the south polar ice at $\approx$ 400 nm [88, 106]. The PMT is coupled with the glass sphere by a clear room temperature vulcanizing silicone gel [105]. The gel-glass combination cuts off short wavelengths at 350 nm which is, however, below the maximum ice transparency. The pressure housing is equipped with a μ-metal grid to shield the magnetic field of the earth to improve the PMT electron collection performance [106]. A flasher board, equipped with 12 individually controllable LEDs (405 nm) can be used for absolute light calibration and ice property investigations [106].

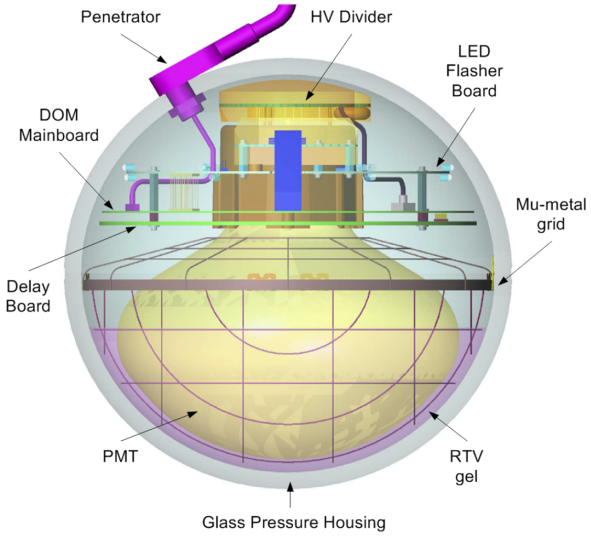


Figure 2.14: IceCube's digital optical module.

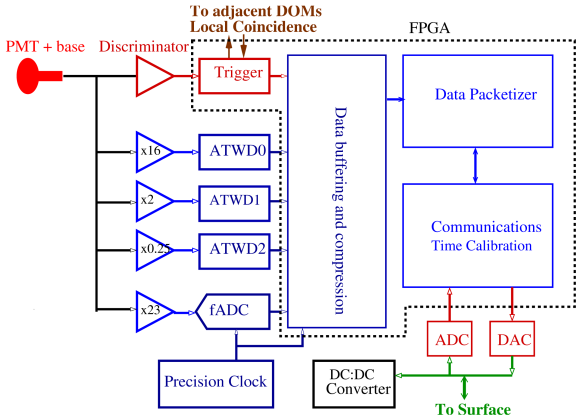


Figure 2.15: DOM electronics block diagram [103].

Figure 2.15 shows a block diagram of the DOM electronics. Digital data processing, buffering, triggering and surface communications are handled in a field-programmable gate array (FPGA) which can be programmed remotely. The digitization and recording of a detected photon waveform from the PMT is triggered by a discriminator if the charge exceeds 0.25 pe. With a delay[17] of 75 ns, the signal is passed to two Analog Transient Waveform Digitizers (ATWD), which record the waveform with a sampling rate of 0.3 GHz [103]. With 128 capacitors, an ATWD can store 420 ns of data. To reduce dead time during digitization of the stored data, the two ATWDs are running in an alternating mode. Additionally, each ATWD chip provides three channels, fed with different amplifications (0.25, 2, 16) to extend the dynamic range. Slower and longer signals can be recorded by a fast analog digital converter (fADC) storing 256 samples at a rate of 40 MHz and thus covering 6.4 µs [105].

[16] The R7081-02 PMTs spectral response spans from 300 nm to 650 nm [106].
[17] Accounting for the delay of the trigger.

The dark photo electron count rate (*dark noise*) at $-40\,°\mathrm{C}$ is $\approx 500\,\mathrm{Hz}$ and is largely caused by radioactive decays in the PMT and pressure housing glass. In addition to the dark noise, further processes can create pulses which deviate from the normal PMT operation. In rare cases of $< 1\,\%$, a photon can pass the photocathode and cause a *pre-pulse*[18] by ejecting a photo electron from a dynode. If the first photo electron is reflected on the first dynode it might trigger the dynode at the second impact and create a *late-pulse* in $\approx 4\,\%$ of the cases and cause a delay of $\approx 60\,\mathrm{ns}$ [106]. Finally, remaining gas molecules can be ionized by the electrons and move towards the previous dynode or photocathode, where a second electron cascade is triggered. These pulses are called *afterpulses*, have a probability of $\approx 6\,\%$ and occur – depending on the ion – with a delay between 300 ns and 1100 ns [106].

While the latter three are connected to external photons, the dark noise appears randomly and isolated. Thus, it can be suppressed by coincidence checks. A full waveform is only stored and transmitted to the surface if the neighboring DOMs on the same string also detected a photon within a time window of $1\,\mu\mathrm{s}$. This mode is known as hard local coincidence (HLC). For more sophisticated noise cleaning methods later in the data processing, reduced hit information is also send for pulses that missed the HLC conditions. This method is called soft local coincidence (SLC) and just stores the time stamp and the charge of the three largest fADC bins [103].

2.5.2 Calibration and data quality

For a later accurate translation of the analog digital converter output into a waveform in mV, the internal voltages need to be calibrated. Additionally, the PMT and electronics delays have to be taken in account. This is done on a monthly basis with the DOM calibration software [107]. The calibration of the high voltage and thus the PMT amplitudes is performed with a 374 nm LED, producing single photon signals in the PMTs. The high voltage is then adjusted to match the desired gain of 10^7, resulting in a charge resolution of 30 % for single photo electrons [103]. PMT transit time, electronic delays, linear behavior and saturation can be evaluated by flashing one or more LEDs on the flasher board.

The synchronization of the DOMs with the surface reference time is realized with the reciprocal active pulsing method (RAPcal) [108]. A precisely timed pulse is send to the DOM which – after a fixed delay – returns the same pulse to the surface. With the same analog digital conversion on both sides an identical transit time can be assumed and the DOM transit time offset is known. This process is repeated every few seconds and achieves a timing accuracy of 1 ns to 2 ns.

The quality of the data and the smooth operation of all detector components is constantly checked by an independent monitoring system. In the course of this thesis, an interactive monitoring web-interface was developed. The system is used regularly by members of the IceCube collaboration organized in a shift plan and the shift takers' feedback

[18] Pre-pulses commonly result in time stamps shifted by $\approx -30\,\mathrm{ns}$ [106].

is valuable input for the evaluation of the data quality[19]. The interface is described in Chapter A of the appendix.

2.5.3 Data acquisition

IceCube's DOMs constantly send digitized, timestamped waveforms to the IceCube Lab at the surface. These multiple data streams are combined and an event within the detector is recorded if certain trigger specifications are fulfilled. For the muon data stream a simple multiplicity trigger (SMT8) is used, requiring the detection of light by 8 different DOMs within 5 µs passing the HLC conditions. In case of a trigger, all waveforms within a time window of ±10 µs are recorded and form one event [103]. For the 79 string configuration of IceCube, this results in an event rate of ≈ 2.2 kHz [109]. Nearly all of these muons originate from atmospheric particle showers created by cosmic ray interactions. Only about one muon within a million events is created by a neutrino of atmospheric origin [109].

From here on, the data is processed event-wise, where one event consists of a collection of raw digitized PMT waveforms. The next step involves a waveform calibration, where the DOM output from the ATWD and fADC is translated into mV and corrected for known electronic effects and possible baseline offsets. Here, the three ATWD channels with different amplifications are combined into one waveform. The channel with the next lower gain is used if a bin is saturated.

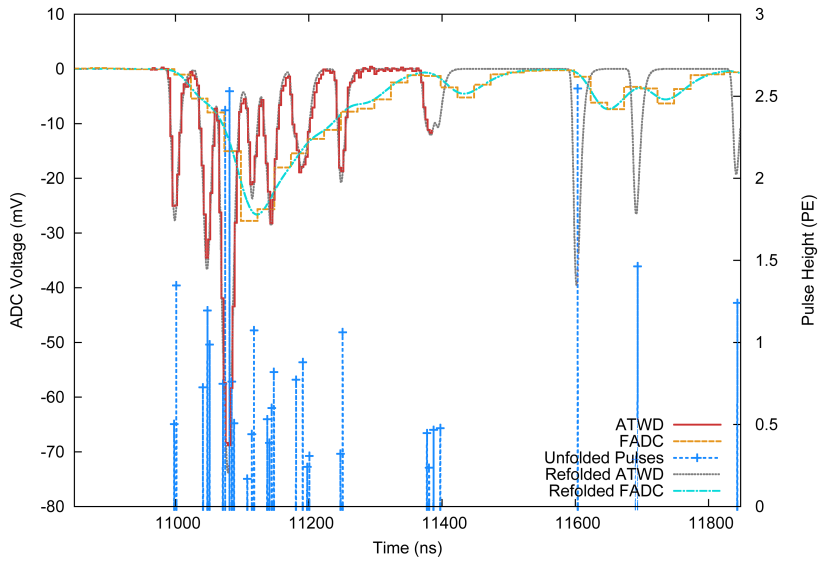


Figure 2.16: Waveform hit extraction: The ATWD output, covering 420 ns, is shown in red and the fADC output is shown in orange. The unfolded photon arrival times are given as blue crosses also indicating the pulse charge. This combination of time and charge is called hit or pulse and is the basis for all track reconstructions. Refolding the incident photons results in the gray and cyan response. Plot modified from [110].

[19]Data taking periods with a problematic detector operation might be excluded in analyses.

The calibrated waveforms could, in principle, already be used for angular track reconstructions as introduced in Section 3.1. However, one needs to define a photon arrival time and a charge from a PMT waveform. In many cases, this is not a trivial task as simple approaches based on thresholds suffer from various problems. If only an over threshold region is considered, some portions of the charge are ignored. Additionally, pile-up effects from consecutive photons – especially important at higher track energies – are handled poorly [110].

These problems are solved by constructing the observed waveform from a linear combination of single photo electron response functions. Details of the unfolding are given in [110] and the process is shown in Figure 2.16. The output is given as blue crosses representing a photon arrival time with a given charge. Large charges are caused by combining two ore more photons. The result is called hit series or pulse series and is the base for all following track reconstructions.

2.5.4 Optical properties of the Antarctic ice

The glacier at the geographic south pole shows a layered structure of optical properties. Down to depths of 1300 m below the surface, the scattering is dominated by air bubbles [111] which are compressed to nonscattering air-hydrate crystals [88, 112] at larger depths. As a result, the photon propagation in the IceCube range between 1450 m and 2450 m below the surface is characterized by thin sub-cm dust layers from ancient volcanic eruptions and broad dust bands caused by climate changes [111].

The ice properties can be examined by observing well defined artificial light sources at different locations. A resolution of a few millimeters can be achieved by slowly lowering a so called dust logger during deployment of IceCube strings, which detects the photon distribution from its own laser beam [111]. The second method, which is used for the ice model described here, quantifies the detector response on light injection from the LED flasher boards of the deployed DOMs (flasher data).

When discussing ice models, it is convenient to combine the scattering length λ_s and the scattering characteristic $\langle \cos\theta \rangle$ into one effective scattering length λ_e [87]

$$\lambda_s \sum_{i=0}^{\infty} \langle \cos\theta \rangle^i \rightarrow \frac{\lambda_s}{1 - \langle \cos\theta \rangle} = \lambda_e \,. \tag{2.33}$$

It expresses the distance that the center of gravity of photons has traveled into the original direction after many scatters. Usually, the reciprocal effective scattering coefficient $b_e = 1/\lambda_e$ and absorption coefficient $a = 1/\lambda_a$ are given instead of λ_e and λ_a.

The ice model currently used in IceCube simulations, Spice Mie, consists of 200 scattering and absorption coefficients in a 10 m binning between 1450 m and 2450 m. These parameters were optimized in a complex process so that the simulated photon propagation matches the flasher data. The resulting ice model properties are shown in 2.17 with $f_{\mathrm{SL}} = 0.45$ and $\langle \cos\theta \rangle = 0.9$. The most prominent feature is the strong dust layer at 2000 m just below the center of the detector. The clearest ice is found below that region where the DeepCore sub-detector is located. Other ice parameters describing the

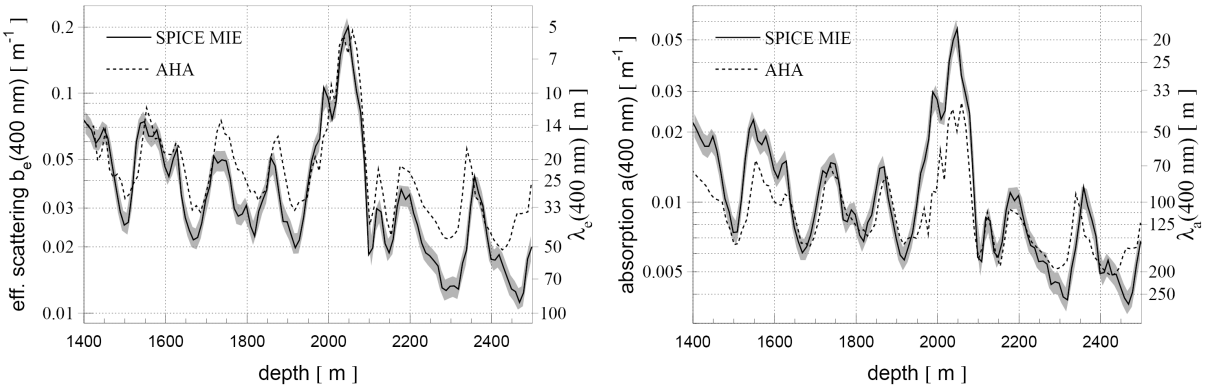


Figure 2.17: Scattering (left) and absorption (right) coefficients are given depending on the depth below the surface for the older AHA ice model [88] and the currently used Spice Mie model [92]. The gray band reflects the estimated uncertainty. Plot from [92].

wavelength dependence of scattering and absorption, or the temperature dependence of absorption are taken from [88]. In the detailed dust logger data, a slight tilt of the ice layers was observed (up to 70 m per km), which is also included in the Spice Mie ice model. Studies show a good agreement between data and Monte Carlo simulation based on Spice Mie [92].

2.5.5 Simulation

The simulation of the complete detection process is fundamental for the performance of a detector. The comparison of simulated and measured data gives valuable information about the understanding of the involved physics, electronic, hardware, and software. In the case of IceCube, this is the propagation and interactions of cosmic rays, atmospheric muons and atmospheric neutrinos, the emission of Cherenkov light, the in-ice light propagation and the complete detector response.

As the Monte Carlo simulation reasonably matches the detector data[20], it is an essential tool to evaluate expected sensitivities (Chapter 5), optimize selection cuts (Chapter 4) or test new methods (Chapter 3). This section describes the IceCube simulation chain and the software used.

2.5.5.1 IceTray - The IceCube software framework

IceTray is the software framework for simulation, processing and analysis in IceCube. The structure represents a chain of independent processing modules, which can be plugged in or changed without altering the rest of the chain. The complete information for an individual event is stored in a frame, which is passed from one module to the next module. For example, the DOM calibration, explained in Section 2.5.3, is implemented in one module which then sends the frame with calibrated waveforms to the next module extracting the photon arrival times, and so on.

[20] A comparison between data and Monte Carlo simulation is shown in Figure B.6 of the appendix.

The first module in the chain usually reads data from a file or a database or, in case of Monte Carlo simulation, it creates a frame by generating a primary particle.

2.5.5.2 Generators

Primary particles are created by so called *generators* in the simulation. Arbitrary neutrino spectra can be injected by the *neutrino-generator* module which is based on the ANIS program [113]. Neutrinos and antineutrinos are created alternating and from all directions at the earth's surface. They are propagated towards the detector undergoing CC and/or NC interactions (see Section 2.4 for cross sections[21]). If the neutrino reaches the detector, it is forced to interact within the generation volume and the resulting cascades and leptons are stored in the frame. The detection probability, production spectrum and the detection volume are combined into one weighting value, called *OneWeight* (OW). Simulated events can later be reweighted according to an arbitrary differential flux function[22] Φ, with the livetime T and the number of generated events n. The resulting weight is

$$w = \Phi \cdot \mathrm{OW} \cdot T/n\,. \tag{2.34}$$

Atmospheric neutrino simulation is obtained for analyses by reweighting the flux according to the predictions given in Figure 2.11 of Section 2.4.4.

Finally, the atmospheric particle showers from cosmic ray interactions are simulated with the *CORSIKA* software package [115]. The fraction of muons that reaches the detector is stored in the frame. The generators described above ultimately create muons and hadronic and electromagnetic cascades. While the light production of the point-like cascades can be directly obtained from parametrizations [86], the muons have to be fully propagated.

2.5.5.3 Muon propagation

Muon propagation in IceCube is performed by the *Muon Monte Carlo (MMC)* module [116], evaluating the energy loss processes covered in Section 2.4.1 in small steps. The primary muon without stochastic energy losses is called *bare muon* and only emits Cherenkov light at the Cherenkov angle. In contrast, the Cherenkov emission from the secondary energy losses (secondaries) along the muon trajectory is smeared around the Cherenkov angle, due to the deviating directions w.r.t. the primary muon [86]. The secondaries with energies above 0.5 GeV are treated individually and parametrizations from [86] provide the light production. The light yield n_l from low-energy stochastic losses below 0.5 GeV can be treated continuously and is given relatively to the Cherenkov light from the bare muon n_b as [86]

$$\mathrm{n_l}(E_\mu) = \mathrm{n_b} \cdot (0.172 + 0.023 \cdot \log(E_\mu[\mathrm{GeV}]))\,, \tag{2.35}$$

[21] Neutrino cross section are currently taken from [80], but are about to be updated from [114].

[22] For example, an $E^{-\alpha}$ power law flux is parametrized by $\Phi = \mathrm{Norm} \cdot (E/\mathrm{GeV})^{-\alpha}$ with a flux scaling Norm in units of $\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

with the muon energy E_μ. The Cherenkov light from the primary muon and secondary cascades is then propagated with a photon propagator.

2.5.5.4 Photon propagation

The Cherenkov photon spectrum from simulated particles and hadronic and electromagnetic cascades is folded with the wavelength dependent acceptance of IceCube DOMs. The different optical properties of the refrozen hole ice, glass, optical gel, and PMT are taken into account. For the primary muon without secondary energy losses, this results in an effective emitted number of 2450 Cherenkov photons per meter for wavelengths in the 300 nm to 600 nm range. All geometrical effects of the DOM-track configuration, like e.g. the angular dependent DOM acceptance, are handled in the propagation.

IceCube's Monte Carlo simulation provides two different approaches for photon propagation. The older Photonics propagation framework [87] can be used to produce tabulated light distributions for all possible source-receiver configurations by propagating many photons on conventional CPUs. These predefined tables can be used in simulations to draw detected photons which heavily speeds up the simulation compared to direct propagation. However, the development of graphical processing units (GPU) enables the possibility of highly parallel computation, perfectly matching the requirements of photon propagation. The Photon Propagation Code (PPC) [117] can directly simulate every single photon capitalizing on the GPU speed boost[23].

In the practice of simulation the straight path s to the next scattering point is drawn from the probability density function $f_{\lambda_s} = \exp(-s/\lambda_s)/\lambda_s$ [87] where the forward deflection angle is sampled from the distribution in Equation 2.27 and the azimuthal angle is drawn from an uniform distribution.

To model the absorption, characterized by the absorption length λ_a, every tracked photon carries its own weight w_{Abs} which is updated during the propagation [87]

$$w_{\text{Abs}} = \exp\left(w_0 - \sum_{i=0}^{n} \frac{\Delta x_i}{\lambda_{a,i}}\right) , \tag{2.36}$$

with the traveled distance Δx_i in a medium with an absorption length $\lambda_{a,i}$. The exact treatment of the weighting now depends on the purpose of the simulation. For direct photon propagation in Monte Carlo simulations the starting weight w_0 can be interpreted as *absorption lengths till absorption* and is sampled from $w_0 = -\ln(\text{rand})$ where rand is a random value between zero and one. The photon disappears when w drops below one while it is fully recorded or detected before. This approach allows a realistic absorption representation and is implemented in the Photon Propagation Code (PPC) [117] and thus called *PPC-absorption*. However, this technique is not suitable for the creation of light distribution tables, because it results in very poor statistics in low light regions far from the track or in dust layers. The photons are absorbed before they reach these regions and thus no information can be recorded here.

[23] On a GPU the propagation task can be executed a factor of ≈ 200 faster than on a CPU.

The solution is the replacement of absorption by a decreasing photon weight. In this case every photon starts with $w_0 = 0$ resulting in a weight $w = 1$. During propagation the weight decreases and photons are recorded with their current weight $w < 1$. The tracking is stopped if a photon leaves the simulation volume or a chosen propagation time is exceeded. This approach is used in the Photonics propagation framework [87] and referred to as *Photonics-absorption*. Contrary to the PPC-absorption, many photons can now reach low light regions – carrying a very low weight – and light timing distributions can be recorded. Both absorption variants give equal distributions for large photon numbers and many light sources.

Accurate photon propagation needs a good description of λ_a and λ_s for every part of the detector. In the layered ice at the south pole the ice parameters can be given as a function of the depth. The most recent collection of ice properties, the Spice[24] Mie ice model, is introduced in Section 2.5.4.

In both cases – the direct GPU propagation and the sampling from predefined tables – the result is a series of photon impacts at IceCube DOMs which are stored in the frame and further processed by the detector simulation.

2.5.5.5 Detector simulation

The first step in the detector simulation is a *NoiseGenerator* module adding the various noise types described in Section 2.5.1. Afterwards, the *DomLauncher* simulates the PMT response waveform, local coincidence checks, the ATWD and fADC digitization and the DOM mainboard. Finally, the trigger conditions are checked by the *trigger-sim* module. From here, the simulation is treated exactly like the real detector data. The next steps are the waveform calibration and the extraction of the photon arrival times as described in the data acquisition Section 2.5.3.

2.5.6 Recent results

The first detection of an high-energy astrophysical neutrino flux is not expected from a single source, but from the integrated diffuse flux from all directions. Given the relatively soft atmospheric background spectra, a diffuse astrophysical neutrino flux is observable by an excess over the predicted background at high energies. Indeed, IceCube found first evidence for a diffuse astrophysical flux at the level of $1.2 \times 10^{-8}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1} \cdot (E/\mathrm{GeV})^{-2}$ [118].

The result was obtained in an analysis based on high-energy neutrino interactions inside the detector. The selection of these events was achieved by an outer veto layer shown in Figure 2.18. Atmospheric muons entering the detector cross the veto regions, deposit light in the veto zone and are rejected. Additionally, high energetic atmospheric neutrinos from the southern hemisphere are accompanied by muons from the same particle shower. Hence, if the veto holds for a downgoing neutrino, it must be of astrophysical origin. The result of the analysis is shown in Figure 2.18. The data in black shows an excess at higher energies

[24] Abbreviation for South Pole Ice

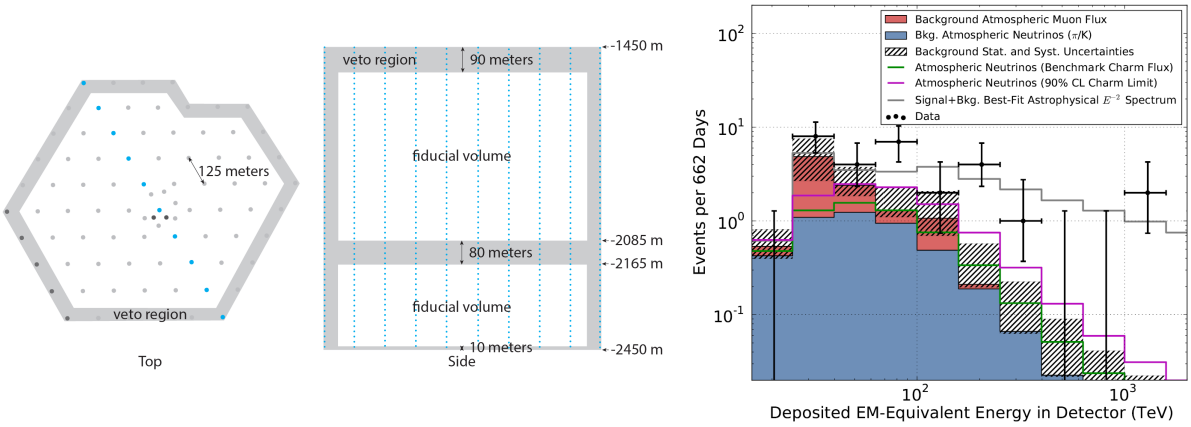


Figure 2.18: Left: The veto region rejects muons entering the detector and is shown in gray. Right: Results from the high-energy starting event analysis [118]. The measured data is shown in black, atmospheric muon predictions in red, conventional atmospheric neutrinos in blue, background uncertainties in black and white stripes and prompt atmospheric neutrinos as well as their 90 % upper limit in green and magenta. The best fit of the background and an astrophysical E^{-2}-flux is given in gray.

and can not be explained by atmospheric background. The background only hypothesis can be rejected at the 4σ level[25].

While the diffuse flux has been observed for the first time, it is still unknown where the neutrinos come from. None of the IceCube point source searches with data from the partial and fully constructed detector identified a single point source [120]. In the following, four approaches with the potential to increase the chances for a point source detection are listed:

- **More statistics:** With a longer exposure and increasing statistics IceCube will improve its point source sensitivity across the years at least by $S/\sqrt{B}$, with the signal statistics S and the background statistics B. However, it is shown in [120] that the increase will even be significantly faster in the next years, as the analysis is not yet background dominated.

- **Time dependent searches:** The atmospheric neutrino background can be reduced by restricting the searches to promising time periods. These can be e.g. periods of higher gamma emission from AGN (flares) or gamma ray bursts. At the time of this writing, no signal was found [51, 121].

- **Stacked searches:** The sensitivity for a flux from a uniform group of potential neutrino emitters can be improved by stacking the data from the directions of the particular candidates. It is shown in [122] that the required flux from one source needed for a discovery is strongly reduced. Stacking analyses fill the gap between the diffuse searches and single point source searches.

[25] Recently, additional data were added to the analyses, increasing the significance to 5.7σ [119].

- **Improved methods:** The discovery potential of IceCube can be enhanced by the development of better methods. These can be e.g. energy and directional reconstructions as well as improved background rejection and signal efficiency.

The latter two approaches are covered in this work in more detail. Chapter 3 introduces a new angular muon reconstruction method with better pointing accuracy while Chapter 4 focuses on signal efficiency and improved background rejection. Finally, in Chapter 5, I describe a stacking analysis, investigating the neutrino flux from blazars.

3

Improved muon track reconstruction

This chapter describes a new method for the muon track reconstruction in IceCube. To start with, Section 3.1 reviews the current reconstruction routines and their weak points. Afterwards, Section 3.2 approaches these issues by discussing a simulation of the muon light distribution in ice based on a precise ice model. The resulting tabulated light distributions are fit by a multidimensional spline surface (Section 3.3) to reduce the file size and to provide a smooth representation. It is then demonstrated in Section 3.4 that the use of these fitted light distributions significantly improves the angular resolution of reconstructed muons. Finally, in Section 3.5 I analyze four physically motivated modifications of the new method which further increase the reconstruction accuracy.

3.1 Muon track reconstruction

The angular reconstruction of muon tracks in IceCube is based on the distribution of detected Cherenkov photons in space and time. By incorporating emission and propagation properties of the light emitted along the straight muon trajectory, the direction of the muon can be evaluated with high precision.

The coordinates used by all reconstructions to define the position and direction of DOM and muon are sketched in Figure 3.1. The trajectory itself is specified by a vertex at the time t_0 and the position $\vec{r}_0$ of a muon with energy E_0. The direction $\vec{p}$ is given in spherical coordinates by the zenith[1] and azimuth angle. A DOM at position $\vec{r}_i$ and perpendicular distance from the track d can be hit by a Cherenkov cone emerging the track according to the Cherenkov angle Θ_C. An unscattered photon arrives at the DOM at the geometrical time t_{geo} (Equation 3.6) with an angle η to the DOM axis.

[1]The zenith angle is defined as $\Theta = 0°$ for a vertically downgoing particle coming from the southern hemisphere, $\Theta = 90°$ for horizontal directions and $\Theta = 180°$ for perfectly upgoing particles from the northern sky.

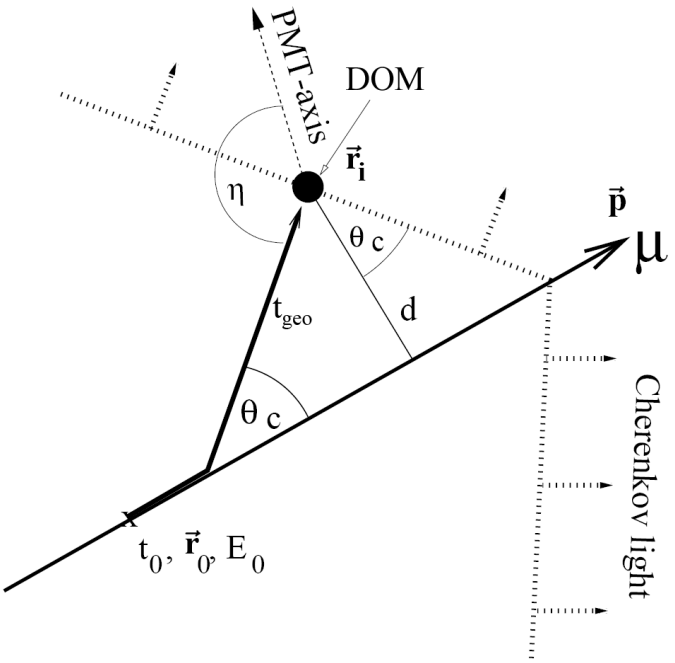


Figure 3.1: Parameters describing the DOM-track configuration [77].

Usually, the reconstruction chain starts with a fast, analytic first guess algorithm like the line-fit which then serves as a starting point (seed) for more sophisticated methods. All the reconstruction methods sketched here are defined in [77] in more detail.

3.1.1 Line-fit

The line-fit, already proposed for the DUMAND II neutrino detector[2], uses a plane wave approximation of the muon light emission and does not rely on a seed track. Optical ice properties and the Cherenkov cone are not taken into account as the light propagation is regarded one-dimensional by projecting the DOM positions onto the muon track. The position of a DOM $\vec{r}_i$ is then given by the track vertex $\vec{r}$, the measured photon arrival time t_i and the light speed along this one-dimensional path $\vec{v}$ as

$$\vec{r}_i \approx \vec{r} + \vec{v} \cdot t_i \quad . \tag{3.1}$$

The most probable track is then evaluated by minimizing the squared sum of the deviation from Equation 3.1

$$\chi^2 = \sum_{i=1}^{N_{\text{hit}}} \left(\vec{r}_i - \vec{r} - \vec{v} \cdot t_i\right)^2 \ , \tag{3.2}$$

[2]DUMAND II was a proposed deep underwater Cherenkov neutrino telescope at Hawaii which was canceled in 1995 due to technical problems.

with the number of hits N_{hit}. The main advantage of the line-fit is that this optimization can be solved analytically quickly, delivering a vertex vector $\vec{r}$ and a velocity vector $\vec{v}$

$$\vec{r} = \langle \vec{r}_i \rangle - \vec{v} \cdot \langle t_i \rangle \quad \text{and} \quad \vec{v} = \frac{\langle \vec{r}_i \cdot t_i \rangle - \langle \vec{r}_i \rangle \cdot \langle t_i \rangle}{\langle t_i^2 \rangle - \langle t_i \rangle^2} \,, \tag{3.3}$$

with $\langle x_i \rangle = {}^1\!/_{N_{\text{hit}}} \sum_i^{N_{\text{hit}}} x_i$ defining the mean of a parameter x. As the determined velocity value depends on the light signature it can be useful in the data selection to separate tracks from cascades.

The resolution of a reconstruction is reflected by the space angle distribution between the simulated Monte Carlo track and the reconstructed track. The median angular resolution of the line-fit is $\approx 5^\circ$ at 10 TeV as shown in Figure 3.3 at data selection level 3 (see Chapter 4).

3.1.2 Likelihood reconstructions

To achieve a better resolution than the line-fit, the light emission direction, its absorption and scattering have to be considered. As a result of respecting the random light propagation, an exact analytical solution does not exist, and the most probable muon track has to be evaluated with maximum likelihood methods. In general, a set of unknown parameters $\vec{\theta}$ can be estimated from a set of measured data points $\vec{x}$ by maximizing the overall likelihood $\mathcal{L}$ with respect to $\vec{\theta}$. In case of statistically independent data points, $\mathcal{L}$ is given by

$$\mathcal{L}\left(\vec{x}|\vec{\theta}\right) = \prod_i p\left(x_i|\vec{\theta}\right) \,, \tag{3.4}$$

where $p\left(x_i|\vec{\theta}\right)$ is the probability to measure data point x_i for a set of parameters $\vec{\theta}$, the so called probability density function (PDF). In the specific context of muon reconstructions in IceCube, $\vec{\theta}$ is the track hypothesis, fully defined by five parameters (degrees of freedom), being the vertex position (x, y, z) and the direction given by a zenith and azimuth angle (Θ, Φ).

For computational reasons, the likelihood is not maximized, but the negative logarithmic likelihood $-\log(\mathcal{L})$ is minimized. This is done with the Simplex algorithm in the Minuit software, included in the ROOT Analysis Framework [123], by varying the track parameters $\vec{\theta}$. All likelihood reconstructions can be performed in multiple iterations with differing starting parameters to avoid converging into local minima.

For angular fits in IceCube, it is convenient not to define the PDF based on absolute photon timestamps but on the delay caused by the scattering. This delay is named time residual t_{res} and can be expressed as the difference between the measured time t_{hit} and the geometrical, unscattered time t_{geo}

$$t_{\text{res}} = t_{\text{hit}} - t_{\text{geo}} \,, \tag{3.5}$$

where t_{geo} is defined by

$$t_{\text{geo}} = t_0 + \frac{|\vec{p} \cdot (\vec{r}_i - \vec{r}_0)| + d \cdot \tan \Theta_C}{c_{\text{vac}}}, \tag{3.6}$$

with the speed of light in vacuum c_{vac}. The PDF can then be expressed as a function $p\left(t_{\text{res}}|\vec{\theta}\right)$ of the time residual and the track hypothesis.

The likelihood definition from Equation 3.4 requires all photons measured in an event's time window to be included. However, in practice, several simplifications are used in IceCube. In the simplest and fastest approach, only the first photon in a DOM is considered as it can be assumed to be scattered least:

$$\mathcal{L}_{\text{SPE}}(\vec{x}|\vec{\theta}) = \prod_{i}^{\text{1st hits}} p(t_{\text{res},i}|\vec{\theta}) \ . \tag{3.7}$$

This reconstruction is called SPE (single-photo-electron) and is usually performed after the line-fit using its result as starting seed. When using the Pandel function defined in Section 3.1.2.1 as PDF, this fit is named PandelSPE and reaches an accuracy of 1° at 10 TeV in Figure 3.3.

The SPE reconstruction completely ignores the total number of photons detected in a DOM. However, the first photon in a DOM that detected a total of n photons would be expected earlier than in a DOM that just measured a single photon. This aspect is considered in the MPE (multi-photo-electron) likelihood description given by

$$\mathcal{L}_{\text{MPE}}(\vec{x}|\vec{\theta}) = \prod_{i}^{\text{1st hits}} n_i \cdot p(t_{\text{res},i}|\vec{\theta}) \cdot (1 - \mathrm{P}(t_{\text{res},i}|\vec{\theta}))^{n_i - 1}, \tag{3.8}$$

$$P(t_{\text{res}}|\vec{\theta}) = \int_{-\infty}^{t_{\text{res}}} p(t|\vec{\theta}) \mathrm{d}t, \tag{3.9}$$

with the cumulative PDF (CDF) $P(t_{\text{res},i}|\vec{\theta})$ and the number of hits in the DOM n_i. Analog to PandelSPE, the MPE likelihood with the Pandel function as PDF is called PandelMPE. For one detected photon per DOM ($n_i = 1$), the MPE likelihood is equal to the SPE variant. As a result, both show similar performance at energies below 10 TeV, while MPE outperforms SPE at higher energies. Figure 3.3 indicates a resolution of 0.6° for PandelMPE compared to 2° for PandelSPE at 10 PeV.

The original likelihood definition from Equation 3.4, using the time information from all detected photons in a DOM, should give the best resolution in theory. However, IceCube studies show a slightly worse performance than MPE, while being slower in computation [124]. The reason is not completely understood but it seems that heavily scattered late photons do not add new information, and, on the contrary, increase systematic uncertainties. So this approach is not used in IceCube and the usual reconstruction chain is line-fit, followed by SPE and MPE.

All likelihood reconstruction methods presented here do not consider PMT noise yet. Thus, a noise hit with no causal connection to the muon track could cause an infinite negative logarithmic likelihood even for the right track hypothesis if the hit's PDF is equal zero at that position. This is solved by adding a small flat noise probability to the PDF. Here a value of 10 Hz is usually used[3].

3.1.2.1 The Pandel function

It is obvious that the performance of the likelihood methods introduced in Section 3.1.2 heavily depends on the quality of the PDF describing the probability of photon arrival times. The shape of this PDF changes with DOM-track parameters like distance and orientation, and thus with the track hypothesis $\vec{\theta}$. For small distances, many unscattered (direct) photons are expected and the PDF peaks at $t_{\text{res}} = 0\,\text{ns}$. With increasing distance, the distribution becomes broader with fewer direct photons.

Studies of the propagation and diffusion of laser light at the Baikal neutrino telescope[4] motivated the use of a gamma function to represent the PDF for photon arrival times. The resulting normalized parametrization is called Pandel function[5] and is given by [125]

$$\begin{aligned} p\left(t_{\text{res}}|\vec{\theta}\right) &= \frac{1}{N\left(d\right)} \frac{\tau^{-(d/\lambda)} \cdot t_{\text{res}}^{(d/\lambda-1)}}{\Gamma\left(d/\lambda\right)} \cdot \exp\left(-t_{\text{res}} \cdot \left(\frac{1}{\tau} + \frac{c_{\text{medium}}}{\lambda_a}\right) - \frac{d}{\lambda_a}\right), \\ N\left(d\right) &= e^{-d/\lambda_a} \cdot \left(1 + \frac{\tau \cdot c_{\text{medium}}}{\lambda_a}\right)^{-d/\lambda}, \end{aligned} \tag{3.10}$$

with the distance d, the scattering length λ, the absorption length λ_a, the Gamma function $\Gamma\left(d/\lambda\right)$ and the speed of light in the given medium c_{medium}. The free parameters λ, λ_a and τ have been adjusted to match results from Monte Carlo simulations:

$$\begin{aligned} \lambda &= 33.3\,\text{m}, \\ \lambda_a &= 98\,\text{m}, \\ \tau &= 557\,\text{ns}. \end{aligned} \tag{3.11}$$

The Pandel function is plotted for three distances in Figure 3.2. An additional modification is necessary for the downwards oriented PMTs in IceCube. Light from a track above a DOM needs to be scattered around the module to hit the PMT. Thus, larger time residuals and a broader PDF are expected which is modeled by replacing the distance d by the larger, angular dependent effective distance d_{eff}

[3] The photomultiplier noise is typically around 500 Hz. However, coincidence conditions reduce it below 10 Hz.

[4] Underwater Cherenkov neutrino telescope in the Baikal lake.

[5] Named after the author Dirk Pandel.

$$
\begin{aligned}
d_{\text{eff}} &= a_0 + a_1 \cdot d\,, \\
a_1 &= 0.84\,, \\
a_0 &= 3.1\,\text{m} - 3.9\,\text{m} \cdot \cos(\eta) + 4.6\,\text{m} \cdot \cos^2\eta\,.
\end{aligned}
\tag{3.12}
$$

The time measurements in IceCube – and thus the time residuals – are not exact but are affected by uncertainties from PMT time information, DOM position and relative DOM clock synchronization. As a result, slightly negative time residuals could occur which by definition have a zero probability in the Pandel function from Equation 3.10. The solution is the convolution with a Gaussian function with a width σ_{PDF} usually chosen to be a few nanoseconds.

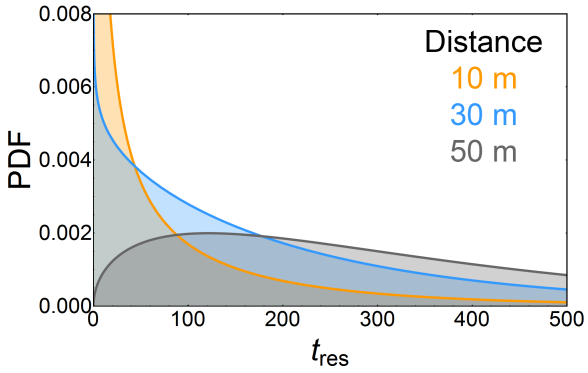


Figure 3.2: The Pandel function at three distances of 10 m, 30 m and 50 m (without Gaussian convolution).

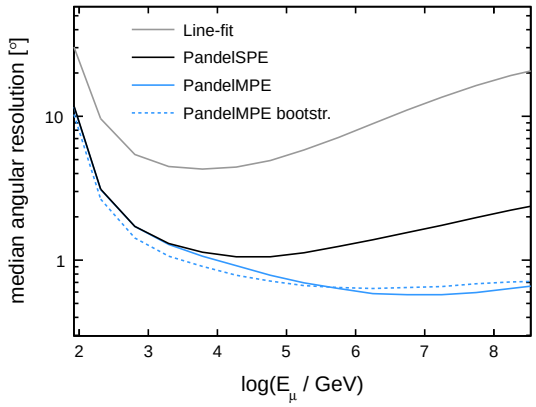


Figure 3.3: IC79 median angular resolution vs. muon energy for line-fit, PandelSPE, PandelMPE and the bootstrapping variant of PandelMPE at data selection level 3 (see Chapter 4).

3.1.3 Bootstrapping

The idea of bootstrapping focuses on a common problem when measuring a sample X_n of n data points from an unknown probability distribution or population F. As there is no mathematical description of F available, it is difficult to tell how good the sample X reflects F. The very simple approach of bootstrapping re-samples $X_n \rightarrow \hat{X}_n$ with replacement to estimate an accuracy by comparing X_n and $\hat{X}_n$ [126].

For IceCube, this technique can in principle be applied to all reconstructions. First, a track is reconstructed using the original series of photon hits resulting in solution 1 (S_1). Next, the original photon hit series is interpreted as a multinominal distribution, where the probability of a particular hit is given by its charge divided by the total charge sum of the event. From this distribution, N new hit series can be re-sampled with replacement.

The result is N sub-samples each with the same total charge as the original hit series but different combinations of the original hits. Each sub-sample is reconstructed independently and the resulting tracks are used as seed for further reconstructions with the original photon series delivering S_2 to S_{1+N}. Finally, S_{N+2} is obtained by seeding an original hit series reconstruction with the average track of the sub-sample reconstructions.

The resulting track is given by the best of the $N+2$ solutions, determined by the lowest negative logarithmic likelihood value. Figure 3.3 shows that bootstrapping outperforms the normal PandelMPE fit below 1 PeV while it performs worse above that energy. The angular resolution of the reconstruction, which was the original purpose of the bootstrapping implementation, can be evaluated from the deviation of the sub-sample tracks. However, it was shown that bootstrapping does not excel established angular resolution estimators introduced in 4.1.2. Hence, bootstrapping is not used for this purpose in IceCube at the moment.

3.1.4 Summary of the discussion

The muon track reconstructions introduced in this section represent a reasonable fit chain from simple first guess techniques to computing intensive algorithms. Especially the PandelMPE reconstruction achieves a sufficient resolution for point source searches. However, an improvement of reconstruction methods would boost the sensitivity of IceCube analyses as the theoretical limitation given by the ν-μ interaction angle (Equation 2.16) is still lower than the angular resolution reached[6].

The primary target for possible improvements is a detailed description of the light distribution PDF which is the centerpiece of all likelihood reconstruction methods. The Pandel function discussed before works surprisingly well being a rather simple representation of such a complex process as light propagation in a natural medium. Nevertheless, it has a few clear disadvantages preventing a better performance. First, auxiliary means like the effective distance d_{eff} are just phenomenological approximations. Second, the gamma function itself does not perfectly fit detailed simulation especially at the region where $d = \lambda$ [77]. But most importantly, it does not consider the depth dependent, layered structure of the south pole ice and the wide range of scattering and absorption lengths measured (see Section 2.4.3 for ice details).

These issues can be solved by the construction of the PDF from tabulated Monte Carlo simulation data for all possible source receiver configurations. In the existing approach based on the Photonics photon propagation package [87], the light distribution is simulated for point-like light sources. In a second step, the light yield from a muon track is constructed through arrangement of point-like sources along the desired trajectory. However, in the end, the resulting PDF did not deliver a superior angular resolution in likelihood reconstructions as the table production method suffers from several technical limitations.

First, the Photonics photon propagation runs on CPUs which limits the achievable table statistics. To avoid large fluctuations, the table binning is chosen rather coarse.

[6]Note that for a point source search in the presence of background, the significance improves linearly with the resolution.

Second, memory limitations also prevent a fine table binning as the point-like light tables already allocate tens of GB in the coarse binning [87] and have to be loaded to the main memory for the construction of the muon light yield. Finally, the combination process for muon tracks involves binning artifacts, further decreasing the quality of the PDF. As a result, the old Pandel function is still used in IceCube's point source analyses up to the IC79 configuration.

One goal of this work is the development of a precise, high-quality PDF for muon reconstruction solving the problems mentioned above, and improving IceCube's pointing accuracy. This process is explained in Section 3.2.

3.2 Light table production: towards a new PDF

Section 3.1.4 focused on the problems of existing light distribution PDFs from Monte Carlo simulations preventing a better angular resolution in IceCube. As a consequence, an improved approach is chosen in this work:

- Direct muon simulation instead of stacking of point-like light sources.
- Photon propagation on graphical processing units (GPU) instead of CPUs to increase production speed and statistics.
- Data storage in very fine-binned tables.
- Table fitting with a multidimensional spline surface to reduce memory usage and provide a smooth fit through statistical fluctuations.

The resulting data has to deliver PDFs for all possible DOM-track configurations in IceCube. The ideal structure is a five-dimensional table representing a cylindrical coordinate system around the muon track which is sketched in Figure 3.4. The muon is defined by a vertex in IceCube's Cartesian coordinates $\hat{x}$, $\hat{y}$ and $\hat{z}$, a zenith angle Θ_S and an azimuth angle Φ_S. The slight tilt of the ice layers (see ice Section 2.5.4) is ignored in this context and the ice is assumed to be isotropic in the $\hat{x}$-$\hat{y}$-plane[7]. As a result, the vertex position in that plane and the azimuth angle are degenerate with respect to light propagation and do not need to be considered in the table structure.

A possible recording point of two Cherenkov photons (blue in Figure 3.4) is shown, defining the storage bin in the table. The depth dependence is described by the z-coordinate Z_S of the point of the closest approach along the muon track. Also the distance ρ is measured from that point. The azimuthal position of the recording point with respect to the muon is given by the angle ϕ. Additionally, the size of the table bins in each dimension is given by $d\rho$, $d\phi$ and dZ respectively. The volume V of the recording cell is indicated by the gray shaded area and is bounded by the table bins or, in the case of

[7]The observed deviation from this isotropy is discussed in [127].

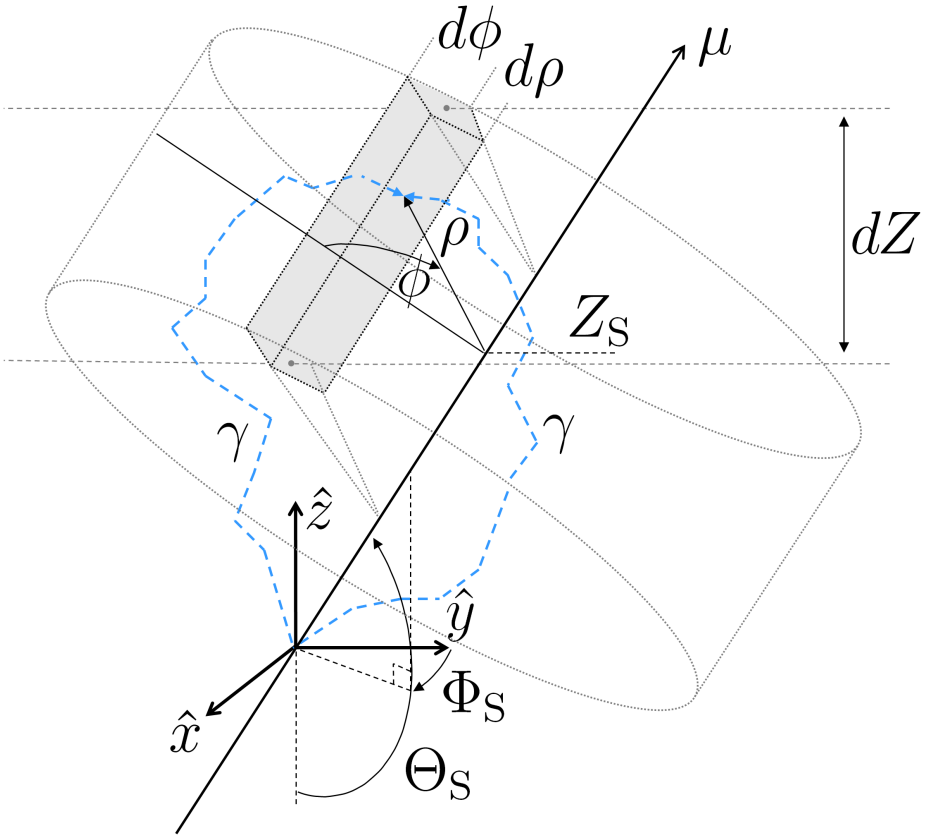


Figure 3.4: Photon propagation coordinate system. Modified from [87].

horizontal tracks, ultimately by the recording box[8]. Finally, a photon's time residual t_{res} is calculated at every recording point with respect to the muon according to Equation 3.5.

The number of bins in each row is summarized in Table 3.1 together with minimum and maximum values and compared to a typical older coarse binning. The time and distance dimensions will show sharp peaks and a rapid change of the distributions at small values, e.g. rather close tracks. Hence, a fine binning is desirable in this case which, however, would be a waste of memory in the diffuse regime far away from the muon and at large time residuals with moderate gradients. As a result, the bin width grows linearly with the bin center value in these two dimensions, while all other dimensions have cells of equal size.

Given the large amount of data expected from the fine binning, the table can not be filled in one process. Instead, it is split by its zenith angle bins. As a side effect, this allows one to create the data of different zenith bins in parallel, speeding up the whole process. With the coordinate system defined, the Monte Carlo production process will be described step by step below.

[8]The recording box has a width of 1200 m and a height of 2000 m around the origin of the IceCube coordinate system in the center of the detector.

Dimension	# bins per dimension		Minimum	Maximum
	old Photonics	this work		
Zenith angle Θ_S	18	18	0°	180°
Depth Z_S	75	75	−850 m	650 m
Azimuthal ϕ	9	36	0°	180°
Distance ρ	30	200	0 m	580 m
Time residual t_{res}	50	210	0 ns	7000 ns

Table 3.1: Photon table binning

In general, the implementation uses IceCube's standard simulation framework introduced in Section 2.5.5. In a first step, 100 TeV muons are injected with a random zenith angle Θ_S within the chosen zenith bin (e.g. 0° to 10° for the first bin) by the simulation module *I3SimpleGenerator*. The particles are released at random positions within a box outside the recording volume. Size and position of the starting box depend on the muon zenith angle and are chosen to satisfy two conditions. First, a minimal distance of 400 m to the virtual detector in its full IC86 configuration needs to be assured to simulate an infinite track in every case. Second, the muons starting in the box should be able to fully enter all ice layers or rather depth bins.

Afterwards, the muon is propagated by the Muon Monte Carlo module *MMC*, thereby loosing energy by radiative processes and ionization as described in Section 2.4.1. According to Equation 2.21, the energy loss per meter as well as the emitted light depend on the current muon energy at the given location. However, to obtain the light yield of a 100 TeV muon along the whole trajectory, the muon is kept at its initial energy at all times for the table production. The effect of this modification on the shape of the PDFs was found to be negligible. This is easy to understand when recalling the small relative, average energy loss of $\approx 3\,\%$ for the 100 TeV-muons after a distance of roughly one photon absorption length of 100 m (Equation 2.21).

In the next step, the *MMC* result is passed to a modified Photon Propagation Code PPC (Section 2.5.5.4) which propagates the Cherenkov light from the primary muon and the secondary energy losses on GPUs. The Spice Mie ice model introduced in Section 2.5.4 provides the ice properties. As PPC was written for direct Monte Carlo production and not for light table creation, major changes had to be implemented. First, to allow the production of light distributions with respect to a certain production process, the light from the primary muon, the quasi continuous secondary energy losses below 0.5 GeV and the stochastic energy losses above 0.5 GeV can now be switched on and off separately (see Sections 2.4.1 and 2.5.5.3 for details on these energy loss processes). Also, the GPU propagation code had to be rewritten to support the Photonics-absorption model necessary for the production of light tables as discussed in Section 2.5.5.4. Additionally, the possibility of recording photons along their path through the ice and not only when hitting IceCube DOMs had to be added. Furthermore, the table structure reflecting the cylindrical coordinate system described above was implemented. Finally, the tilt of ice layers was switched

off as it cannot be reflected in the table structure chosen and the memory management was modified to be able to efficiently write thousands of recording points from the GPU to the table in the main memory instead of a few DOM hits per muon.

The emitted photons are recorded every Δs meters[9] on their way through the ice. At each recording point, the current photon weight is calculated according to

$$w_\gamma = w_{\text{Abs}} \cdot w_{\text{Eff}}(\vec{p}) \,, \tag{3.13}$$

with the current photon absorption weight w_{Abs} according to Equation 2.36 and the angular sensitivity w_{Eff}[10] of IceCube DOMs depending on the photon direction $\vec{p}$ at the current position. The table coordinates distance ρ, azimuthal angle ϕ, depth Z_{S}, and time residual t_{res}, are calculated with respect to the emitting muon and the photon weight is stored in the corresponding bin.

The expected number of photo electrons n_p from a DOM in a table cell is evaluated according to

$$n_p = \frac{\Delta s \cdot A_{\text{DOM}}}{V_{\text{Total}}} \sum_{k=1}^{n_\gamma} w_\gamma(k) \,, \tag{3.14}$$

with the distance between recording points Δs, the DOM area A_{DOM}, the total cell recording volume V_{Total}, and the sum of all n_γ photon weights $w_\gamma(k)$ recorded in the current bin. V_{Total} is the sum of the bin volumes[11] (shaded area in 3.4) from all simulated muons

$$V_{\text{Total}} = \sum_{i=1}^{n_\mu} V_{\text{Bin}}(i) \,. \tag{3.15}$$

To obtain sufficiently low statistical uncertainties in all table bins a total of 10 000 muons were simulated for each zenith bin, which takes ≈ 2 days on one NVIDIA GeForce GTX 480 GPU. It was found that the time needed for coping the numerous recording points from the GPU memory to the main memory limited the propagation calculation speed, resulting in a GPU load of $\approx 50\,\%$. Hence, two separate processes could be run on one GPU with alternating copy operations to fully utilize the computation potential. As a result, the full 18 zenith bins were produced in less than a week on four GPUs.

3.3 Table-fitting with Photospline

The light distribution tables produced in the course of this work and described in Section 3.2 represent the time residual PDF for all possible source-receiver combinations with

[9] For the light of just the primary muon, a value of $\Delta s = 5\,\text{m}$ was used and $\Delta s = 40\,\text{m}$ for high energetic stochastic losses.

[10] w_{Eff} is modeled by a parametrization that also respects the different structure of the melted and refrozen bore hole ice.

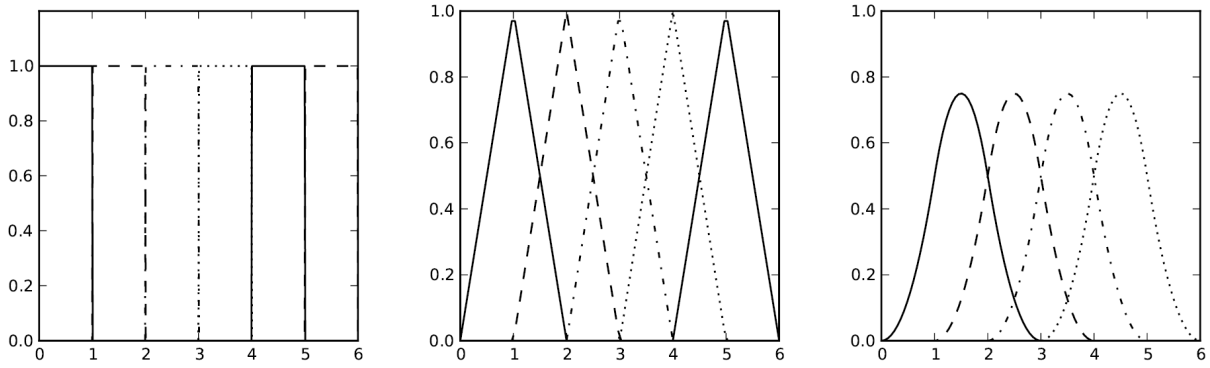
[11] Volume calculation code courtesy of Bastian Terlinde.

roughly 2 billion entries. In principle they contain all information needed for the likelihood reconstructions defined in Section 3.1.2. However, the total size of the tables is about 20 GB. The memory configuration of current high performance computing systems usually does not allow the use of such large data structures. The possible solution of partially reading the currently needed section of the table from the hard disk during the reconstruction process would result in a very poor performance. An additional problem could arise from fluctuations in lower statistic regions of the table, affecting the likelihood minimization performance. Finally, the tables don't consider detector timing uncertainties and should be convolved with a Gauss function for reconstruction.

In summary, a highly compressed representation of the table, that smooths statistical fluctuations and is foldable with a Gauss function, is desired. These features are provided by the Photospline software [128] which is designed to fit large multidimensional tables with a multidimensional spline surface.

3.3.1 Photospline

Splines are piecewise defined polynomial functions frequently used in interpolation and fitting problems. The piecewise definition avoids problems like divergences and oscillations occurring e.g. in high order polynomials. In Photospline, the table is fitted by so called basis splines (B-splines) of order n, which can be determined and evaluated quickly and have $n-1$ continuous derivatives. They are formed by iterative self convolution as demonstrated in Figure 3.5 [110]. The fitted surface is a linear combination of B-splines, each defined by a knot vector $\vec{k}$.



(a) Order-0 B-splines (b) Order-1 B-splines (c) Order-2 B-splines

Figure 3.5: Example for B-splines of order $n = 0$, $n = 1$ and $n = 2$ from [110]. Folding two B-splines of order 0, results in a B-spline of order 1, and so on.

When recalling the shapes of the time residual PDF (for instance, see Figure 3.2), it is obvious that the peak value spans several orders of magnitude, mainly depending on the distance. Thus it is more convenient to fit the cumulative distribution (CDF) which varies between 0 and 1. The PDF can be obtained from the derivative at any point. The

monotonicity of the CDF in t_{res} is forced by the non-negative least squares fitting method, as a decreasing CDF would result in a negative PDF. An adjustable smoothness parameter λ is used in the B-spline fitting process to penalize rapid changes in curvature, thereby ensuring a fit that smooths over statistical fluctuations. Finally, the resulting spline surface can be convolved with an Gauss function in the time dimension to account for the various time resolution uncertainties (Fig. 3.9). A detailed conceptual and mathematical description of Photospline is given in [128].

3.3.2 Fitting procedure

Several steps are necessary to prepare the table data for the fitting process. It is convenient to separate the timing information represented by normalized PDFs from the expected light yield per table cell. Both are fitted independently and the PDFs are converted to CDFs for the fitting. Due to the large range of the expected light, the data are fit in a logarithmic scale.

One spline surface covers one zenith angle-depth combination. For the binning shown in Table 3.1, this results in $18 \times 75 = 1350$ three-dimensional surfaces which are fit independently and connected by pseudo-interpolation afterwards [128].

Important tuning parameters are the number of spline knots and the knot spacing in the knot vector $\vec{k}$. Corresponding to the size of the table bins, it is reasonable to choose a denser configuration where rapid gradient changes are expected. Thus, the knot positions in the time and distance dimension are chosen equidistant on a logarithmic scale while for the azimuthal dimension a linear spacing was utilized. The number of knots was kept as low as possible while maintaining a sufficiently accurate fit result. The configuration finally chosen is summarized in Table 3.2.

Dimension	**# knots**	**Spacing**
Azimuthal ϕ	25	linear
Distance ρ	40	exponential
Time residual t_{res}	30	exponential

Table 3.2: Photospline knot configuration used in this work.

All dimensions are fit with B-splines of order $n = 2$ except of the time CDF dimension. The time dimension has to be derived to obtain the PDF which reduces the spline order by one. Figure 3.5 shows that at least order-2 splines are necessary for a smooth fit. Thus, for a PDF represented by order-2 B-splines, the CDF has to be fit with $n = 3$. A reasonable smoothing was achieved with $\lambda = 0.01$ and the time dimension was forced to increase monotonously. Figure 3.6 shows a typical CDF fit result (lines) to the table data shown as crosses for different DOM-track distances. A data-fit comparison for PDFs is plotted in Figure 3.7. A geometric effect of the downward orientated PMTs is clearly visible. For large azimuthal angles, the light has to be scattered around the DOM to be

detected, resulting in longer time residuals. In general, a very robust and accurate CDF fit was achieved.

The separate fit of the expected number of photo electrons is drawn in Figure 3.8. The plot nicely shows the effect of the ice layers by the wave structure of the red line representing a DOM below the track undergoing different absorption lengths. In the case of a DOM in the same ice layer (green), one constant absorptivity creates a perfect exponential decline. The blue line enters the big dust layer after 120 m, causing a rapid light decrease. The given numbers show the expected light from the primary *bare* muon and do not reflect secondary energy losses or effects like cable shadowing from IceCube strings. Additionally, a second set of splines has been produced, representing the light yield from high-energy stochastic losses (used in Section 3.5).

The light distribution tables produced in Section 3.2 have been successfully fit with Photospline. The result is a smooth representation of all CDFs and PDFs in one file with a size of 206 MB and all expected light yields in one file with a size of 7 MB. Compared to the total table size of 20 GB, the memory load was reduced by a factor of ≈ 100. Section 3.4 investigates the reconstruction performance with this new PDF.

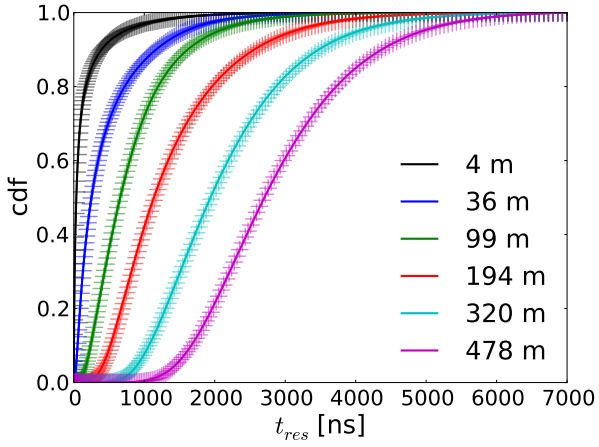


Figure 3.6: CDF spline fit (lines) for muons with Θ_S between 50° and 60°, Z_S between −270 m and −290 m and ϕ between 175° and 180° for different distances ρ. Corresponding table data is shown as crosses.

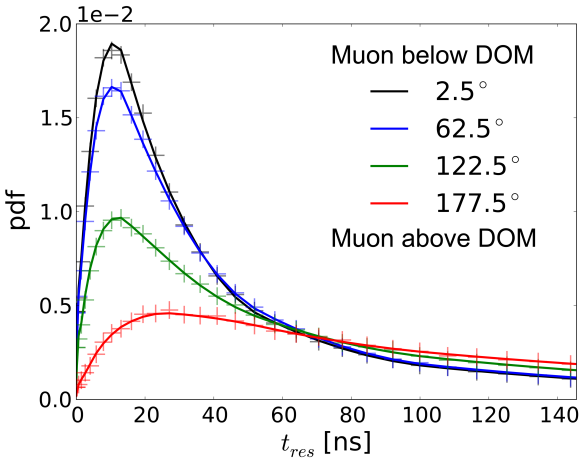


Figure 3.7: PDF spline derived from CDF fit (lines) for muons with Θ_S between 50° and 60°, Z_S between −270 m and −290 m and ρ between 23.7 m and 24.8 m for different azimuthal angles ϕ. Corresponding table data is shown as crosses (error bars are too small to show).

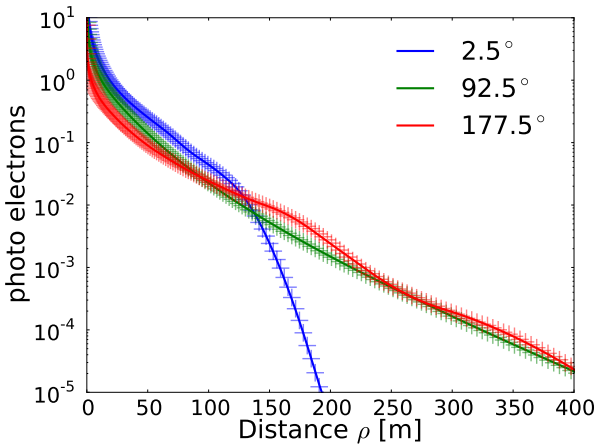


Figure 3.8: Expected light yield fit (lines) for muons with Θ_S between 50° and 60° and Z_S between −230 m and −250 m for different azimuthal angles ϕ. Corresponding table data is shown as crosses.

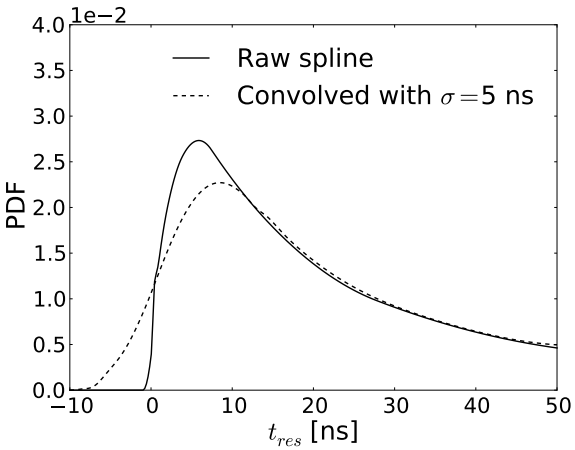


Figure 3.9: Raw spline result and convolved variant with $\sigma_{PDF} = 5$ ns

3.4 Spline reconstruction results

One advantage of the spline representation described in Section 3.3 is that it fits the CDF and derives the PDF from it. Hence, all information needed for IceCube's current best likelihood reconstruction, the MPE fit (Equation 3.8), is already available. To evaluate the accuracy, a new reconstruction module called *spline-reco* was developed[12]. It replaces the Pandel function with the spline PDF and the cumulative Pandel function with the spline CDF. The minimization routine and noise handling (10 Hz) stay unchanged. For the PDF convolution, the standard value for PandelMPE $\sigma_{PDF} = 4$ ns is used. Figure 3.10 compares the new method called SplineMPE (orange) with the existing routines. A significant improvement in angular resolution is achieved through the whole energy range relevant for IceCube. The best median resolution of 0.4° is reached at the highest energies while the improvement is smaller at low energies. This SplineMPE reconstruction with the splines produced in the course of this work is the currently best angular muon fit in IceCube. It is used in all point source analyses with the full IC86 detector [120]. Compared to PandelMPE, the computation is slower by a factor of 2.7. In Section 3.5 we investigate whether this SplineMPE reconstruction can be further improved by additional modifications.

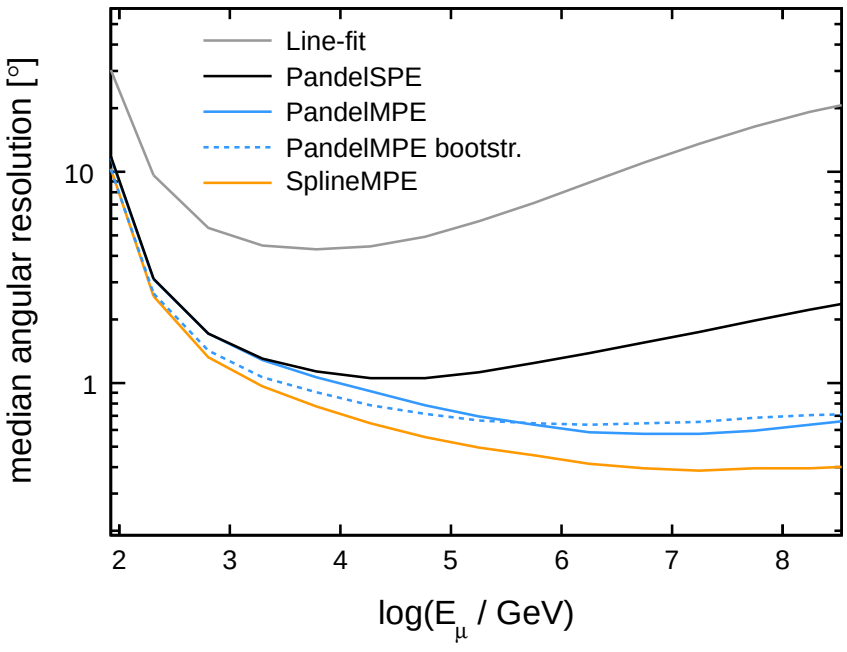


Figure 3.10: IC79 median angular resolution vs. muon energy for line-fit, PandelSPE, PandelMPE, the bootstrapping variant of PandelMPE and SplineMPE at data selection level 3 (see Chapter 4).

[12]First version by Jacob Feintzeig.

3.5 Further likelihood-fit improvements

The performance of likelihood reconstruction methods, introduced in Section 3.1, depends heavily on the quality of the underlying PDFs. The probability distribution of photon arrival times at a DOM for a certain muon track has to be as close to reality as possible. The replacement of the workhorse Pandel function, assuming homogeneous ice, by new spline PDFs, created with a precise, layered ice model (Section 3.2) significantly improves the angular resolution.

Nevertheless, the time delay distribution is still solely based on the Cherenkov emission from the primary muon, also referred to as *bare muon*. Thus, the PDF does not contain aspects like secondary stochastic energy losses (also named secondaries or cascades), accurate PMT noise modeling and the correct treatment of detector uncertainties. This chapter addresses these issues and investigates whether the reconstruction based on the hypothesis of a continuous, averaged light emission from an infinite track can be further improved.

For detailed resolution analyses it is convenient to examine the full information contained in the actual point spread function instead of the median angular resolution. The point spread function is built by the normalized distribution of the space angles, or delta angle, between true Monte Carlo tracks and reconstructions and is shown on the left in Figure 3.11 for SplineMPE and PandelMPE for an E^{-2} neutrino spectrum. The cumulative point spread function is shown in the middle. The relative cumulative improvement of the SplineMPE with respect to the PandelMPE is obtained by dividing the cumulative SplineMPE distribution by the cumulative PandelMPE distributions. The result is shown in the right plot in Figure 3.11. Here it is easy to see that SplineMPE reconstructs 10 % more events closer than 1° and 20 % more events closer than 0.4° to the true Monte Carlo track compared to PandelMPE. The statistical uncertainties, hardly visible in Figure 3.11, have been derived by splitting the total event sample into 10 sub-samples and evaluating the standard deviation per bin.

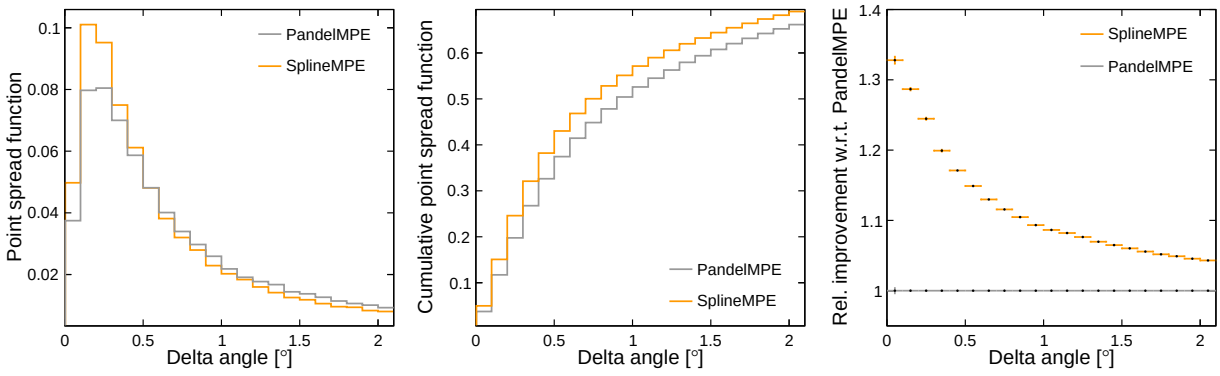


Figure 3.11: Point spread function (left), cumulative point spread function (middle) and relative improvement of the cumulative point spread functions with respect to PandelMPE (right) for an E^{-2} neutrino spectrum.

This representation will be used in the course of this section to measure the impact of several modifications. In contrast to Figure 3.11, the results will not be shown relative to PandelMPE but relative to the best current muon reconstruction in IceCube, which is SplineMPE. The SplineMPE reference is set up as in Section 3.4 with a PDF convolution assuming $\sigma_{\mathrm{PDF}} = 4\,\mathrm{ns}$ to account for time uncertainties and a flat noise probability of 10 Hz to handle PMT noise.

3.5.1 Hit cleaning with the Kolmogorov-Smirnov test

As stated above, there are several light emission processes that are not yet covered by the spline PDF. For example, large stochastic energy losses along high-energy muon tracks do not fit the PDFs assuming a continuous emission. Figure 3.12 demonstrates how a 0.1 TeV cascade early along the muon track adds late hits compared to a continuous emission. At a first glance this might not affect the MPE likelihood from Equation 3.8, as only the time of the first photon is considered. However, the total DOM charge n defines how narrow the likelihood function is. Thus, using a PDF which is modeling a continuous emission alongside with the actual DOM charge also including secondaries can cause too narrow MPE likelihood distributions with peaks shifted towards lower time residuals. This might lead to a worse angular reconstruction. In other words, the PDF does not fit the real light distribution.

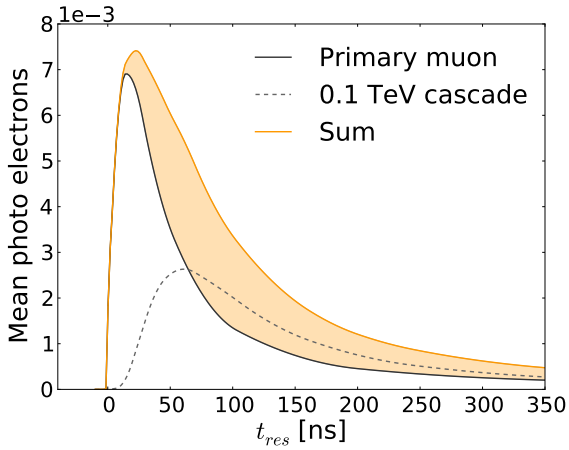


Figure 3.12: The light from a 0.1 TeV cascade (dashed) early on the muon track travels slower than the muon in ice. As a result, DOMs along the muon track detect additional late hits from this stochastic energy loss. The orange shaded area indicates the additional photons compared to the bare muon (solid).

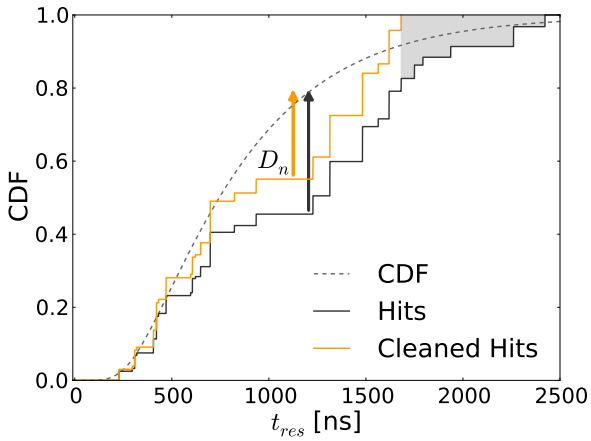


Figure 3.13: The maximal difference D_n between the CDF $P(t_{res})$ (dashed gray) and the empirical distribution function of the hits $F_n(t_{res})$ (black) is called Kolmogorov-Smirnov statistic (black arrow). In the case of additional late hits, it can be reduced (orange arrow) by discarding these late hits (shaded area) until the cleaned sample (orange) passes the Kolmogorov-Smirnov test.

The Kolmogorov-Smirnov test quantifies whether a statistic sample follows a given distribution. It could help to clean an event's pulse series from these late hits caused by stochastic energy losses by comparing it with the CDF of a bare muon. The Kolmogorov-Smirnov statistic D_n and the empirical distribution function $F_n(t_{res})$ shown in Figure 3.13 are given by

$$D_n = \sup_{t_{res}} |F_n(t_{\mathrm{res}}) - P\left(t_{\mathrm{res}}\right)| \quad \text{with} \tag{3.16}$$

$$F_n(t_{\mathrm{res}}) = \frac{1}{\sum_{i=1}^{n} q_i} \sum_{i=1}^{n} I_{\mathrm{t}_{\mathrm{res}_i} \leq t_{res}} \cdot q_i \quad . \tag{3.17}$$

Here t_{res} is the time residual, sup the supremum, $P\left(t_{\mathrm{res}}\right)$ the cumulative probability density function, n the number of hits, $I_{\mathrm{t}_{\mathrm{res}_i} \leq t_{res}}$ the indicator function[13], and q_i the charge of hit number i. The hit sample passes the Kolmogorov-Smirnov test if D_n is smaller than the critical value $K_{\alpha,n}$, depending on the sample size n and the desired significance level α. These critical values are tabulated in Table B.1 up to a sample size of 20 hits and can be estimated for bigger hit series by [129]

$$K_{\alpha,n} = \sqrt{\frac{\ln(\frac{2}{\alpha})}{2n}} \quad . \tag{3.18}$$

In the first step of the reconstruction, the measured photon distribution $F_n(t_{res})$ is compared to the CDF $P\left(t_{\mathrm{res}}\right)$ based on the reconstruction starting parameters (seed). If the Kolmogorov-Smirnov test for a DOM fails because there are too many late hits, indicated by $F_n(t_{\mathrm{res}}) - P\left(t_{\mathrm{res}}\right) < 0$, the last pulse is discarded. This procedure repeats until the hit series satisfies the Kolmogorov-Smirnov criterion or until there is just one pulse left. The method is sketched in Figure 3.13 and was implemented by Bastian Terlinde. Afterwards, the usual likelihood minimization is performed with every DOM's Kolmogorov-Smirnov approved charge. It was tested to repeat the Kolmogorov-Smirnov test from time to time in the likelihood optimization with updated parameters, but the changing likelihood landscape leads to minimization problems. It is obvious that the Kolmogorov-Smirnov test, being only performed once at the beginning, needs a good starting track to perform reasonably. In this study, a previous MPE reconstruction using the Pandel function was used as a seed.

Figure 3.14 shows the impact of the Kolmogorov-Smirnov cleaning on the reconstruction performance for E^{-1} and E^{-2} spectra. For E^{-2} there's a small relative improvement as only about 2 % more events are reconstructed closer as 0.2° to the true track. For the E^{-1} spectrum the effect is even smaller. In both cases, the result does not depend strongly on the significance level but overall $\alpha = 20\,\%$ gives the best result. As expected, there is no significant change for low energy events in E^{-3} (Figure B.1 of the appendix) where most DOMs rarely detect more than one photon anyway. One has to conclude that the

[13]Here the indicator function I is equal to unity if the time residual of pulse number i is smaller or equal t_{res} and 0 otherwise.

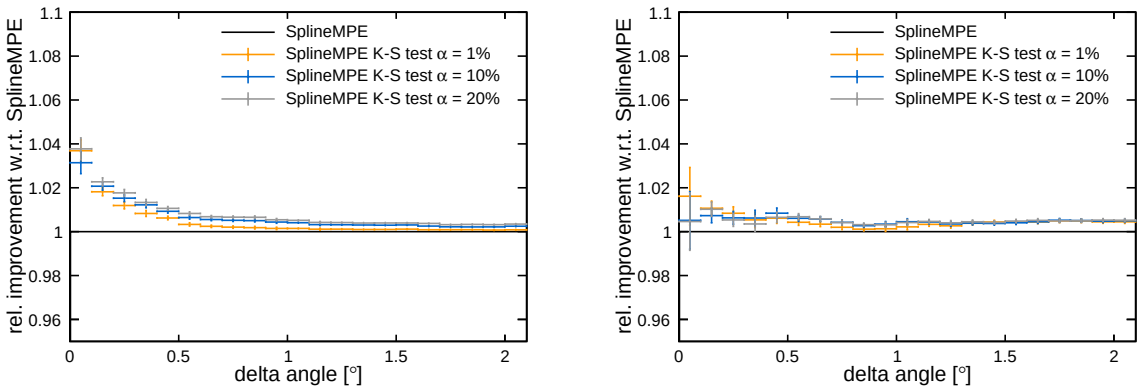


Figure 3.14: Reconstruction with Kolmogorov-Smirnov cleaned pulses: The plots show the relative improvement of the cumulative point spread function with respect to the default SplineMPE for an E^{-2} spectrum on the left and an E^{-1} spectrum on the right for three different significance levels α.

influence of stochastic processes is rather weak. However, the method might improve if a better starting track is given, which is analyzed in Section 3.5.5.

3.5.2 Including mean stochastic energy losses

The spline PDF used in the reconstruction so far reflects the Cherenkov light distribution of a bare neutrino induced primary muon. However, with rising muon energy, further light production processes like ionization, pair production, bremsstrahlung, and nuclear interactions become increasingly dominant. These hadronic or electromagnetic cascades not only emit light exactly at but also around the muon Cherenkov angle as displayed in Figure 3.15. This leads to differently shaped PDFs as visible in Figure 3.16. It has to be noted that the high-energy secondaries are not emitted continuously but they are randomly distributed along the muon track, which can not exactly be reflected by the model of a continuous, averaged light emission used here. Nevertheless, combining the different shaped PDFs weighted by their light contribution depending on the muon energy might result in a more realistic overall PDF and improve the reconstruction performance. The resulting PDF should have the shape of the bare muon PDF at low energies and be blended into the stochastics PDF with rising energy. For this purpose the PDF and absolute number of expected photo electrons n of the stochastic energy losses must be known.

As introduced in Section 2.5.5.3, in practice the light emitting effects are arranged in three groups:

- Primary bare muon Cherenkov emission (b),
- Cherenkov emission from stochastic energy losses above 0.5 GeV (h) and
- Cherenkov emission from stochastic energy losses below 0.5 GeV (l).

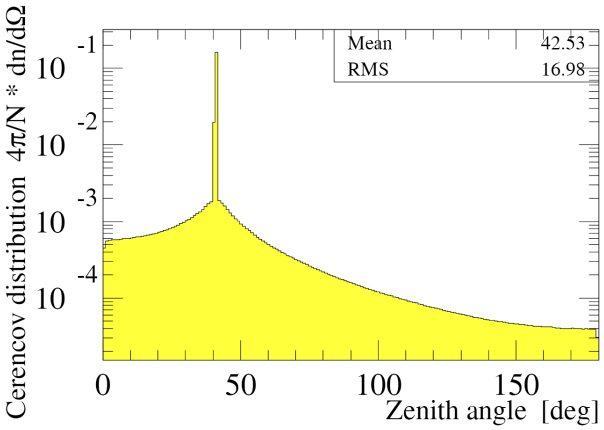


Figure 3.15: Cherenkov distribution for 300 GeV muons including only secondaries below 0.5 GeV simulated with GEANT from [86].

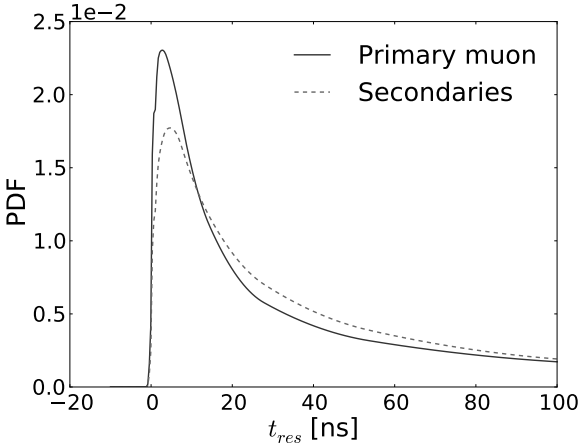


Figure 3.16: PDFs of bare muon and averaged secondaries only: The light emission of stochastic energy losses along the muon track is smeared around the Cherenkov angle, which causes a slightly broader averaged PDF compared to the radiation from the primary muon.

The probability density function of the bare muon PDF p_b and the expected number of photo electrons of the bare muon n_b are available through the splines described in Section 3.2. For secondaries above 0.5 GeV another set of splines has been produced accordingly, but with bare muon light and low energy secondaries switched off. This gives access to the average high-energy secondary PDF p_h and the average expected number of photo electrons $n_\mathrm{h}(E_\mu)$ of high energetic cascades along the muon track. According to Section 4.1.1, $n_\mathrm{h}(E_\mu)$ scales linearly with the muon energy E_μ.

Below 0.5 GeV, the emission from stochastic energy losses is assumed to be continuous and the expected number of photo electrons n_l of low energy secondaries can be parametrized depending on the muon energy E_μ and the primary muon light n_b by Equation 2.35. The total expected number of photo electrons $n_\mathrm{tot}(E_\mu)$, calculated by adding up the three contributions yielding

$$n_\mathrm{tot}(E_\mu) = n_\mathrm{h}(E_\mu) + n_\mathrm{l}(E_\mu) + n_\mathrm{b}\,, \tag{3.19}$$

is plotted in Figure 3.17. Below 1 TeV, the expected light yield mainly consists of the energy independent Cherenkov radiation from the primary muon n_b with a slight energy dependence from low energy secondaries $n_\mathrm{l}(E_\mu)$, while it is dominated by high energetic stochastic energy losses $n_\mathrm{h}(E_\mu)$ above 1 TeV.

The angular distribution of Cherenkov light from [86] is considered to be identical for low and high energetic secondaries in IceCube's energy range, and so too the PDFs are assumed to be identical:

$$p_{\mathrm{h}} = p_{\mathrm{l}} \quad . \tag{3.20}$$

With the information above, the overall PDF can be constructed as a weighted average

$$p_{\mathrm{tot}}(E_\mu) = \frac{(n_{\mathrm{h}}(E_\mu) + n_{\mathrm{l}}(E_\mu)) \cdot p_{\mathrm{h}} + n_{\mathrm{b}} \cdot p_{\mathrm{b}}}{n_{\mathrm{tot}}(E_\mu)} \,, \tag{3.21}$$

where E_μ can be estimated with one of IceCube's muon energy estimators introduced in Section 4.1.1.

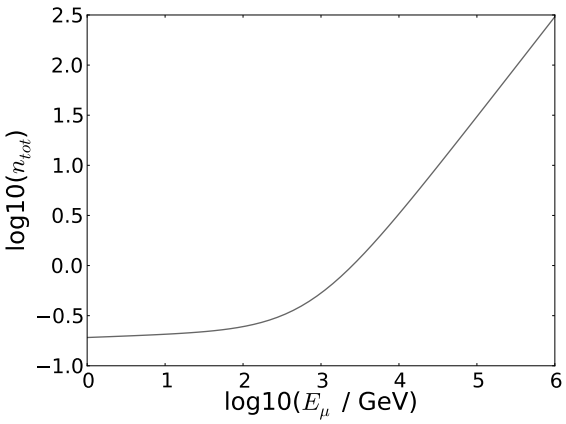


Figure 3.17: Expected number of photo electrons $n_{\mathrm{tot}}(E_\mu)$ from a vertical downgoing muon detected by a DOM at a distance of 50 m depending on the muon energy.

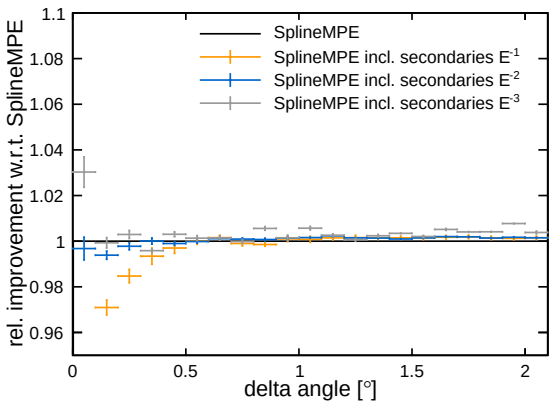


Figure 3.18: Relative difference of the cumulative point spread function of the reconstruction including averaged stochastic energy losses compared to the default SplineMPE for three different spectra.

Using this merged PDF p_{tot} in a MPE likelihood reconstruction gives the results shown in Figure 3.18. No improvement in resolution could be achieved. Especially regarding the worse performance for high-energy events in the E^{-1} weighting, one has to conclude that the highly stochastic nature of the secondary energy losses can not be approximated by the averaged PDFs in this approach.

3.5.3 Accurate noise modeling

To handle PMT noise hits in reconstruction, usually a flat dark noise probability of 10 Hz is added to the likelihood. The flatness reflects the noise distribution of uncleaned hits well as shown in 3.19, where the noise hits form a flat shape between the time residuals -3000 ns and 3000 ns. The slopes outside this range are due to varying event time window positions. However, the reconstructions in IceCube rely on cleaned hit series, for which Figure 3.20 displays a rather different situation. As cleaning algorithms basically discard isolated hits,

the noise is under control very well in the realm where it is unlikely to see photons from the muon track, which is the case for very negative time residuals for example. However, the coincidence criteria cause a clustering of noise around photon hits from the track which cannot be represented by a flat noise term. The use of a PDF that is reflecting the actual noise probabilities more closely might improve the track reconstruction.

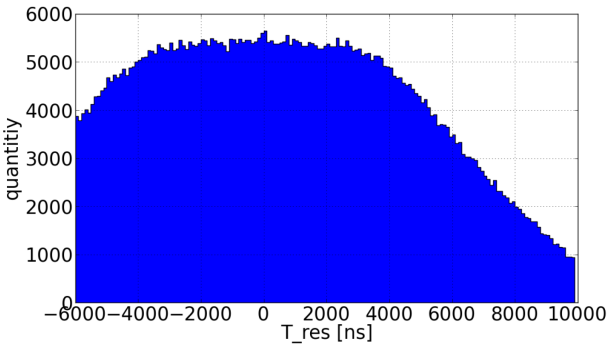


Figure 3.19: Distribution of uncleaned noise hits in IceCube [130].

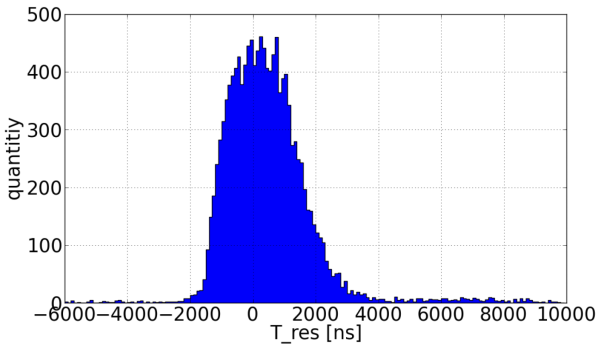


Figure 3.20: Distribution of cleaned noise hits in IceCube [130].

The shape of all different noise contributions, being dark PMT noise, pre/late pulses and after pulses, has been studied in [130] for different hit cleanings. Instead of a flat noise probability, these exact noise distributions could be used in reconstructions. To merge these exact noise terms with the actual spline PDF, their contributions to the total hit count in the current DOM must be known. This weight mainly depends on the expected number of photo electrons from the track as a higher photon count from the track allows a higher rate of dark noise to pass the cleaning. Also, pre/late and after pulses are directly correlated to photon hits. The contribution w from a certain hit type in a DOM is defined by the average number of photo electrons per hit DOM and is shown in Figure 3.21 depending on the expected number of photo electrons n_{tot} calculated after Equation 3.19. The contribution of photons originating from a muon track is shown in blue (physics pulses), dark PMT noise in red, after pulses in green and pre/late pulses in yellow. The left side of the plot represent regions far away from the track, where almost every hit DOM just contains one dark noise hit (red). With a rising expected number of photo electrons n_{tot}, more and more DOMs are hit by photons from the muon track (physics, blue) which lets the dark noise ratio per hit DOM decrease. In the region above $n_{\text{tot}} = 1$, all DOMs on average see light from the track and all dark noise hits pass the cleaning, which makes w for dark noise stay stable before the plot runs out of statistics. Pre/late and after pulses basically follow the photon hits from the muon track.

By studying the noise distributions in [130], the shapes of the different noise contributions have been modeled and merged with the muon PDF using a parametrization of the weight w shown in Figure 3.21 depending on the total expected number of photo electrons $n_{\text{tot}}(E_\mu)$.

The result of the reconstruction using this more exact noise model is plotted in Figure 3.22. There are about 2 % more events closer than 0.5° to the true track for an E^{-3}

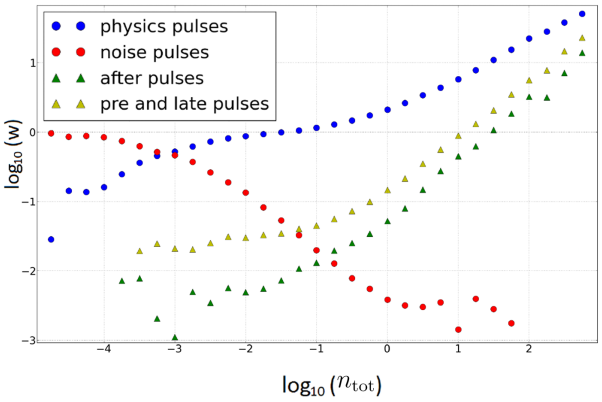


Figure 3.21: Number of hits per hit DOM w for different hit types depending on the expected number of photo electrons n_{tot} calculated after Equation 3.19. Plot taken from [130].

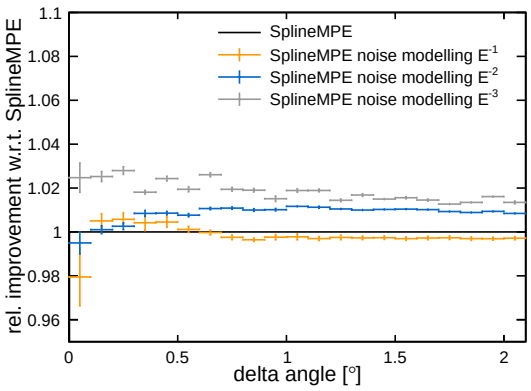


Figure 3.22: Relative difference of the cumulative point spread function of the reconstruction with an accurate noise model compared to the default SplineMPE for three different spectra.

spectrum. For harder spectra the impact is weaker, because single noise hits loose importance with much light in the detector. As a conclusion, one can state that a very slight improvement is possible.

3.5.4 MPE likelihood convolution

The time measurement in IceCube is affected by various uncertainties listed in Table 3.3. The absolute GPS timing and the relative DOM timing depend on the clock synchronization accuracy and the data acquisition involves uncertainties from signal transmission times [108]. The PMT time measurement varies with the electron transit time spread [106] and the exact DOM position in the bore hole involves a geometrical uncertainty constrained by the hole diameter of 60 cm. Finally, an additional component arises from the frequency dependent speed of light in ice [131]. These effects are respected in the track reconstruction by convolving the PDF in the MPE likelihood by a Gaussian. Best results are achieved with $\sigma_{\text{PDF}} = 4\,\text{ns}$ so far.

However, a closer look into the contributions to the overall timing precision reveals that only the PMT timing uncertainty affects every hit separately and should be applied to the PDF. The other effects[14] influence every hit in the same way or rather the whole DOM and should be adopted to the whole MPE likelihood function from Equation 3.8. In other words, not only the PDF has to be folded by a Gauss function but also the MPE likelihood. To test this approach, a new MPE likelihood convolution stage defined by σ_{LLH} is needed besides the already used PDF convolution with σ_{PDF}.

[14]The absolute GPS timing is irrelevant for the reconstruction, though, as it affects all DOMs in the same manner.

Source	Uncertainty
abs. GPS timing	10 ns per 24 hours [108]
rel. DOM timing	1 ns to 2 ns [108]
PMT	2 ns [106]
Data acquisition	0.7 ns [108]
Geometry	1 ns [103]
Dispersion	1.5 ns after 40 m [131]

Table 3.3: Uncertainties in the IceCube time measurement.

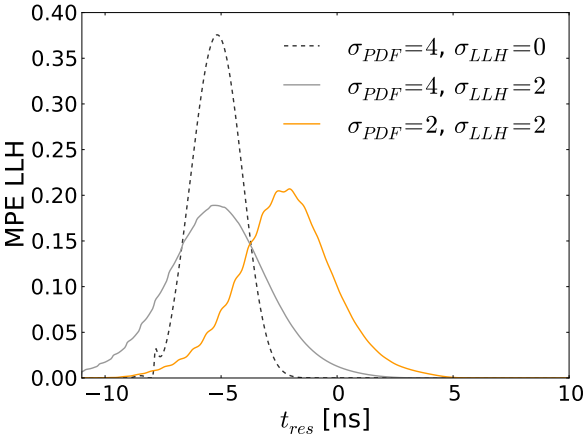


Figure 3.23: MPE likelihood for three different convolution configurations.

With respect to reconstruction speed, a fast recursive approximation algorithm for Gaussian convolution [132] implemented by [133] has been added to the existing reconstruction framework. The MPE likelihood is calculated at several sample points around the requested time residual based on the convolved PDF, which is accounting for PMT uncertainties. From these sample points, the smeared MPE likelihood is calculated with adjustable accuracy depending on sample point density and recursion step count[15].

This MPE convolution solves two problems of the old methods especially appearing with high DOM charge n. Without applying σ_{LLH}, the MPE likelihood can be more narrow than IceCube's timing accuracy as $(1 - P(t_{\text{res}}))$ is raised by $n - 1$ in the MPE definition from Equation 3.8. Secondly, by applying a σ_{PDF} that is too big, the MPE likelihood will be unphysically shifted towards lower time residuals. Both behaviors are displayed in Figure 3.23 as well as the ideal approach. The default configuration without σ_{LLH} applied will form unphysical sharp peaks for high DOM charges (dashed), which can by solved by convolution with a likelihood smearing $\sigma_{\text{LLH}} = 2\,\text{ns}$ (solid gray). Applying a realistic σ_{PDF} and σ_{LLH} of 2 ns each shifts the MPE likelihood towards positive time residuals (orange).

In order to find the best setup, the PDF convolution has been set to $\sigma_{\text{PDF}} = 2\,\text{ns}$ to simulate the PMT timing uncertainty. Figure 3.24[16] shows the relative improvement of different σ_{LLH} values compared to the previous default setup using $\sigma_{\text{PDF}} = 4\,\text{ns}$ and $\sigma_{\text{LLH}} = 0\,\text{ns}$. The optimal setting are determined to be $\sigma_{\text{PDF}} \approx 2\,\text{ns}$ and $\sigma_{\text{LLH}} \approx 2\,\text{ns}$ which reconstructs about 5 % more muon tracks closer than 0.2° to the true track.

This improvement comes with a decreased reconstruction speed by a factor of 1.5 mainly caused by the multiple calculation of the MPE likelihood at the sample points. The execution time per event is not multiplied by the number of sample points (20), as repetitive spline lookups with the same DOM-track configuration are much faster.

[15] 20 sample points, distributed equidistantly over a range of $6 \cdot \sigma_{\text{LLH}}$ work reasonable with 4 recursion steps.

[16] The plot for an E^{-3} spectrum is shown in Figure B.2 of the appendix.

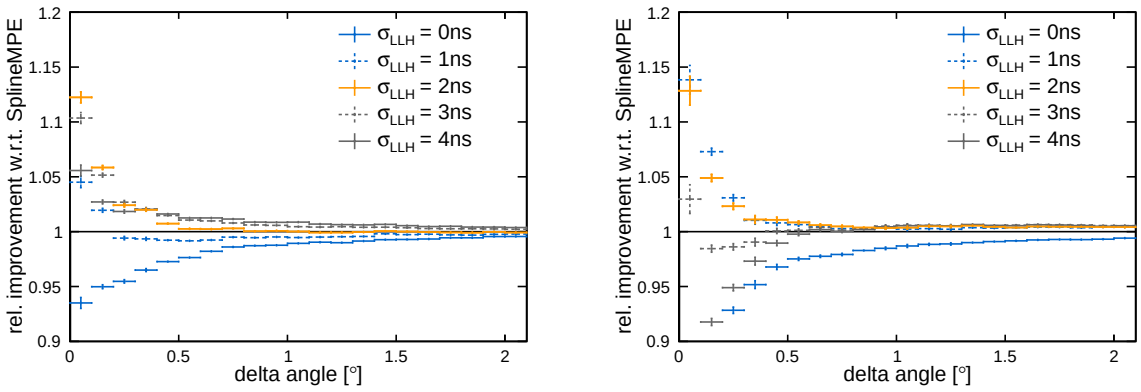


Figure 3.24: Relative improvement of the cumulative point spread function with respect to the default SplineMPE reconstruction with $\sigma_{\mathrm{PDF}} = 4\,\mathrm{ns}$ and $\sigma_{\mathrm{LLH}} = 0\,\mathrm{ns}$. Different σ_{LLH} values are shown for an E^{-2} (left) and E^{-1} (right) neutrino spectrum, while σ_{PDF} is fixed to 2 ns.

3.5.5 Combination

In this section, the modifications examined above are combined into one modified SplineMPE reconstruction. As the inclusion of the mean stochastic losses spline did not lead to an improvement it is disregarded in this attempt. The best configuration from the separate examinations are applied in the combination: The Kolmogorov-Smirnov test is set up with a significance level $\alpha = 20\,\%$, the MPE likelihood convolution works with $\sigma_{\mathrm{PDF}} = 2\,\mathrm{ns}$ and $\sigma_{\mathrm{LLH}} = 2\,\mathrm{ns}$ and the noise modeling is enabled for the input of cleaned photon hits. The PandelMPE reconstruction will serve as starting track (seed). The point has been made, that the Kolmogorov-Smirnov test depends on a good starting track. This motivates the addition of a second subsequent modified SplineMPE reconstruction with a refined seed from the first run.

The top plot in Figure 3.25 shows the relative improvement of the cumulative point spread function with respect to the default SplineMPE reconstruction for an E^{-1}, E^{-2} and E^{-3} neutrino spectrum in orange, blue and gray, respectively. The first run, seeded with PandelMPE, is shown as dashed lines. The second reconstruction, seeded with the result of the first run, is shown as solid lines and yields an additional improvement by the refined Kolmogorov-Smirnov test.

For an E^{-1} spectrum, an additional 10 % of the events are reconstructed closer than 0.3° to the true track. The same improvement was achieved for E^{-2} at 0.2°. For a soft E^{-3} spectrum the gain at 0.2° is 5 %. The lower plot gives the median angular resolution as function of the neutrino energy for the PandelMPE in blue, the default SplineMPE in orange and the second run of the modified SplineMPE in red. As already suggested in the upper plot, a better accuracy is realized through the whole energy range by the modifications with an impact rising with energy. The average interaction angle between muon and neutrino from Equation 2.16 is shown for convenience. The small differences to Figures 3.3 and 3.10 are due to a slightly different data selection level. Overall, the

observed accuracy gain is better than the sum of the separate modifications. It has to be explained by a synergetic effect of the three different contributions, resulting in a more realistic likelihood description.

Note that at lower energies, the minimization sometimes does not find the optimal minimum, which is evident from seeding the reconstruction with the true value known only in the Monte Carlo. It has been shown [134] that thereby the SplineMPE reconstruction further improves the angular resolution below 10 TeV, when performed in up to 16 iterations with randomized starting parameters.

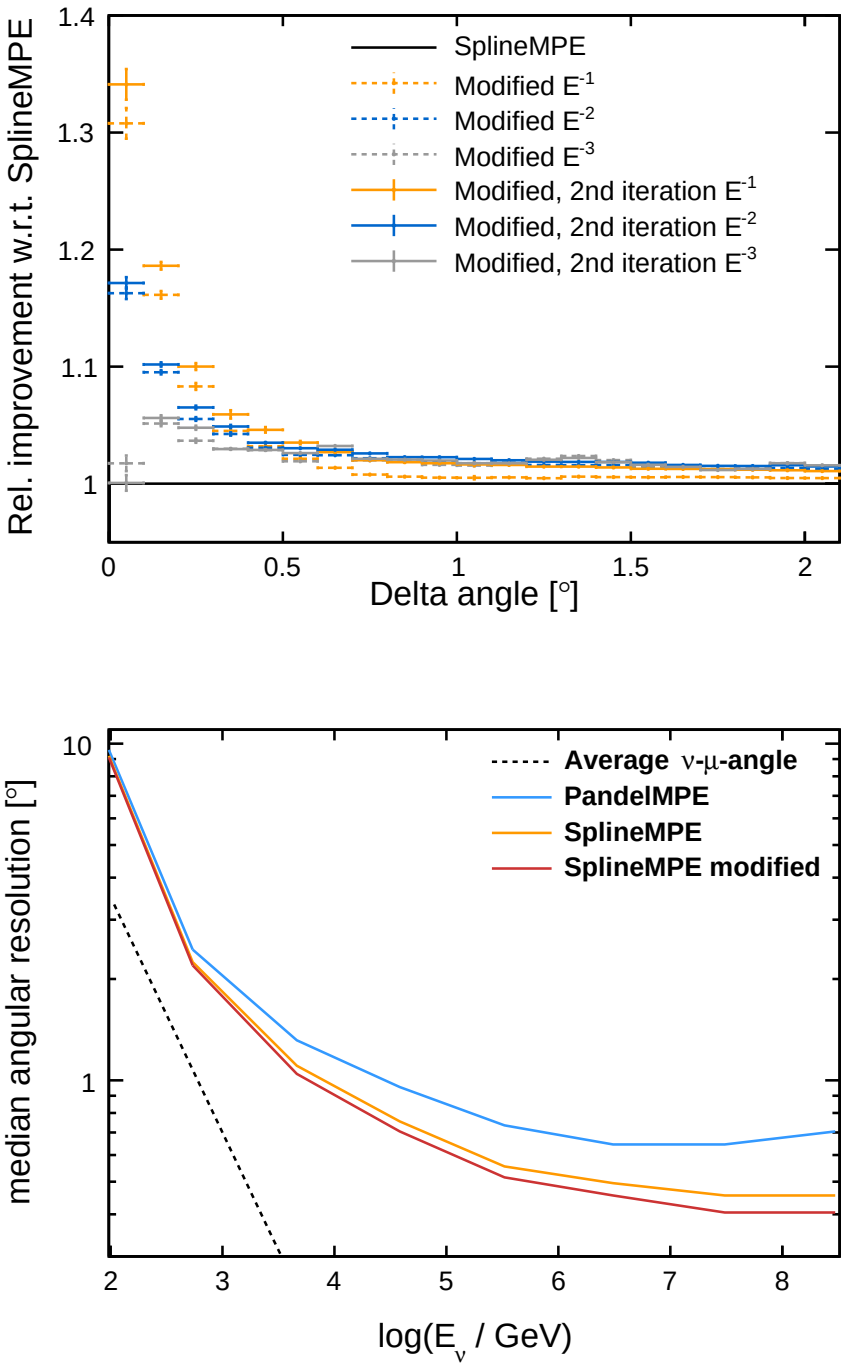


Figure 3.25: Upper Plot: Relative improvement of the cumulative point spread function of the SplineMPE incorporating three modifications, with respect to the default SplineMPE reconstruction for three different neutrino spectra. The dashed lines represent the first run seeded with PandelMPE while the second reconstruction is seeded with the result of the first run and shown solid. Lower Plot: Median angular resolution depending on the neutrino energy for the PandelMPE in blue, the default SplineMPE in orange and the second run of the modified SplineMPE in red. The average interaction angle between muon and neutrino from Equation 2.16 is shown dashed as a reference.

3.6 Summary

This chapter described the road towards a significantly improved angular muon reconstruction in IceCube. Very fine binned Cherenkov light distributions were produced on graphical processing units for all muon configurations. These tables were fit by a multi-dimensional spline surface to reduce the memory requirement and to smooth statistical fluctuations. In addition, detector timing uncertainties can be considered by convolution with a Gauss function. The splines can be used as a probability density function in likelihood reconstructions replacing older, less detailed parametrizations like the Pandel function. This results in a better median angular resolution through the neutrino energy range from 0.1 TeV to 100 PeV, reaching 0.45° at the highest energies compared to 0.7° from previous methods. The method was first available for the analysis of the data from the first year of the fully constructed IC86 detector. As it currently delivers the best angular resolution it was established as default reconstruction in IceCube point source analyses since then [120].

Additionally, it was shown here that three modifications in the likelihood further improve the accuracy. These changes reflect an accurate PMT noise modeling, a photon hit cleaning with the Kolmogorov-Smirnov test and the correct treatment of timing uncertainties. Again, the resolution is improved through the whole energy range above 0.1 TeV down to 0.4° at the highest energies.

As these numbers were evaluated on an intermediate data selection level, a better resolution is expected at the final analysis level covered in Chapter 4, favoring well reconstructed and well contained tracks.

4

Data filtering and event selection

The muon reconstruction method developed in Chapter 3 was used for the data of the fully constructed detector in the IC86 configuration. Unfortunately, it was not yet available, when the data of the 79 string configuration of IceCube was processed and analyzed [109]. This chapter describes the reprocessing of that data to include the SplineMPE reconstruction. There are several reasons supporting this plan. First, the improved pointing should lead to a better discovery potential for point source searches. Second, the 79 string configuration is very close to the complete detector. That means that with high event counts and good angular resolution the impact of this data set in combination with other detector configurations is significant. Third, a uniform data set is preferable. Finally, as there already exists a data set based on the old reconstruction approach, this is a good opportunity to precisely study the changes and improvements coming along with the new method by direct comparisons.

To begin with, Section 4.1 introduces the cut and quality parameters used in the selection process. To add the new reconstruction, it is not necessary to start all over from the output of IceCube's data acquisition, described in Section 2.5.3. Instead, the reprocessing, described in Section 4.2.3, is performed on an intermediate selection level, introduced in Sections 4.2.1 and 4.2.2. Once the SplineMPE reconstruction is added, the last selection step is based on boosted decision trees as specified in Section 4.3. During this process, there will be frequent comparisons with the old sample based on the PandelMPE reconstruction from [109]. The final result, named IC79b, is analyzed in Section 4.4 and is used in the blazar stacking analysis described in Chapter 5.

4.1 Cut and quality parameters for the analysis

A wide range of variables is used to distinguish between an astrophysical signal and atmospheric background. These reflect different properties of each event and are described in

this section. The main focus is on the energy estimators, the angular resolution estimator and the SplineMPE reconstruction (Chapter 3), which are the main components of the point source analyses introduced in Section 5.2.

4.1.1 Energy estimators

The estimated value of the muon energy E_μ is used as a cut parameter in the data reduction process, distinguishing the hard astrophysical neutrino flux and the soft atmospheric background. As most high-energy muons are not fully contained in the detector, the original neutrino energy can not be reconstructed. However, the muon energy E_μ can still be evaluated from the energy loss along its trajectory. Equation 2.21 shows the approximately linear connection between E_μ and the energy loss per meter dE_μ/dx above a muon energy of 1 TeV. In addition, the expected number of photons detected by a PMT is directly proportional to dE_μ/dx [81]. So E_μ can be evaluated by matching the detector response by linear scaling of a template track with either a certain dE_μ/dx or E_μ. The template function $\Lambda(\vec{\theta})$, providing the expected number of photons detected by a PMT for all DOM-track configurations $\vec{\theta}$, is approximated analytically or taken from lookup tables produced in Monte Carlo simulation.

Monte Carlo approach: In the Monte Carlo approach, $\Lambda(\vec{\theta})$ was evaluated beforehand from simulation and the data is stored in tables structured like those described in Chapter 3. Here, the template reflects a muon with an average energy loss per meter of $dE_\mu/dx = 1\,\mathrm{GeV\,m^{-1}}$. These tables already contain the expected detector response as well as the position-dependent properties of the photon propagation through the layered ice [135]. Due to the linearity described above, the event's dE_μ/dx can easily be calculated by comparing the number of detected photons k in all DOMs with the number of expected photons $\Lambda(\vec{\theta})$ from the template:

$$dE_\mu/dx = \left(\sum_{\mathrm{DOMs}} k_j\right) / \left(\sum_{\mathrm{DOMs}} \Lambda(\vec{\theta})_j\right) \cdot 1\,\mathrm{GeV\,m^{-1}}\,. \tag{4.1}$$

A correction function evaluated from Monte Carlo events then translates the obtained dE_μ/dx into the muon energy E_μ. The Truncated Energy (TE) method introduced in Section 4.1.1.3 uses a variation of this process.

Analytical approach: Estimating the expected number of photons analytically clearly cannot represent the complex light propagation through the layered ice; however, it results in a good computational performance. Thus, such fast methods can be applied early in the processing chain when the data rate is still high.

For the hypothesis of an infinite muon track, Cherenkov radiation is emitted at the Cherenkov angle and decreases exponentially with distance due to absorption. Close to the track within one scattering length, the photon density is proportional to $1/r$ where r is the distance of the closest approach [135]. Further away, the light is diffused and

the photon density is given by $\exp(-r/\lambda_p)/\sqrt{r}$ assuming a random walk. Here, λ_p is the characteristic propagation length defined as $\lambda_p = \sqrt{\lambda_a \lambda_e/3}$, where λ_a is the absorption length and λ_e is the effective scattering length, defined in Section 2.5.4 [88][135]. The two distance regimes are then combined by an empirical functional form, matching both cases. The number of photons is given by [135]

$$\Lambda(r) = l_0 A \cdot \frac{1}{2\pi \sin\Theta_c} e^{-r/\lambda_p} \frac{1}{\sqrt{\lambda_\mu r} \tanh\sqrt{r/\lambda_\mu}}\,, \tag{4.2}$$

$$\text{with} \quad \sqrt{\lambda_\mu} = \frac{\lambda_c}{\sin\Theta_c}\sqrt{\frac{2}{\pi\lambda_p}} \quad \text{and} \quad \lambda_c = \frac{\lambda_e}{3} e^{\lambda_e/\lambda_a}\,. \tag{4.3}$$

Here, l_0 is the number of photons emitted per meter along the track, A is the effective collection area of the DOM and Θ_c is the Cherenkov angle. In this case, contrary to the Monte Carlo approach, $\Lambda(r)$ reflects the expected average number of detected photons from a template track with a reference muon energy for the given DOM-track configuration. Again because of the linear connection, the expected number of photons from the muon can be expressed directly with the muon energy by $E_\mu \cdot \Lambda(r)$. The likelihood to detect k photons, given an expectation of $E_\mu \cdot \Lambda(r)$ photons plus ρ noise photons and a muon energy E_μ, follows a Poisson distribution [135]

$$\mathcal{L} = \frac{(E_\mu \Lambda(r) + \rho)^k}{k!} \cdot e^{-(E_\mu \Lambda(r)+\rho)}\,, \tag{4.4}$$

$$\ln\mathcal{L} = k\ln(E_\mu\Lambda(r)+\rho) - (E_\mu\Lambda(r)+\rho) - \ln(k!)\,. \tag{4.5}$$

Maximizing the likelihood sum $\sum \ln\mathcal{L}(E_\mu)$ of all DOMs delivers the most probable muon energy for the given light pattern. If this is not possible analytically, the $-\ln\mathcal{L}$ is optimized with numerical minimization routines. MuE and MuEX specified in Sections 4.1.1.1 and 4.1.1.2 apply this method.

4.1.1.1 Analytical implementation

The first implementation of the analytic approach described above is called MuE and gives a moderate energy resolution shown in Figure 4.1 compared to MuEX and TE (Sections 4.1.1.2 and 4.1.1.3). The reason is that the likelihood description in Equation 4.5 expects a precise model of the expected light yield which clearly cannot be delivered by the approximation from Equation 4.2. But even with very accurate Monte Carlo tables, the model still describes an averaged continuous emission and cannot account for high energetic stochastic losses like e.g. Bremsstrahlung occurring along the muon track. So the result is very often biased to match the charge detected by DOMs close to such an energy loss. This is visible in MuE's larger high-energy tail in Figure 4.1. The advanced estimators MuEX and TE approach that issue in different ways.

However, the fact that MuE is very sensitive to stochastic losses turns it into a tool to distinguish between atmospheric muon bundles and single high energetic muons of possibly

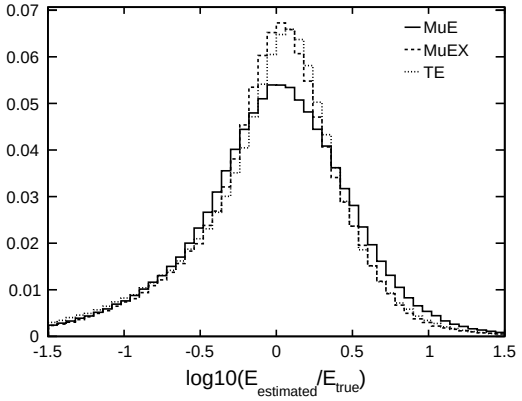


Figure 4.1: The energy resolution is defined as the difference of the logarithms of the estimated energy and the true energy and is presented for MuE (solid), MuEX (dashed) and TE (dots). Fitting a Gaussian function (not shown) results in logarithmic deviations of $\sigma_{\mathrm{MuE}} = 0.48$, $\sigma_{\mathrm{MuEX}} = 0.41$ and $\sigma_{\mathrm{TE}} = 0.43$.

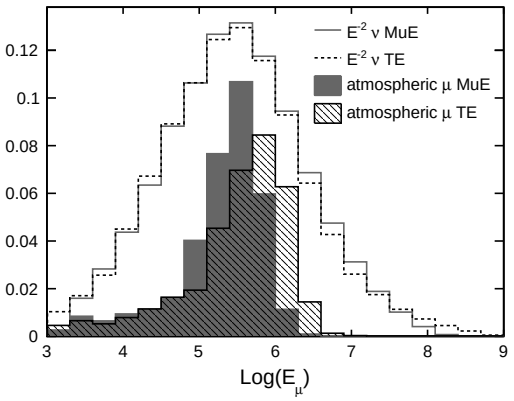


Figure 4.2: The plot shows the muon energy estimation for atmospheric muons (shaded) and E^{-2} neutrinos (lines) from TE (black) and MuE (gray). Only tracks from the southern sky have been plotted (zenith angle <85°).

astrophysical origin. A bundle of a few parallel low energetic muons from one air shower can produce the same amount of light as one high-energy event but with a much smoother emission profile. Figure 4.2 shows that MuE and TE have a similar energy distribution for the astrophysical neutrinos simulating an E^{-2} spectrum (dashed and solid lines). But for the higher energy part of the atmospheric muons (shaded areas), where bundles dominate over single muons, MuE gives lower energy estimations. This is helpful for likelihood analyses utilizing the energy for signal-background discrimination in the southern sky as described in Sections 4.4 and 5.2.

4.1.1.2 Improved analytical implementation

Several changes have been applied to MuE resulting in an improved energy estimator called MuEX. Some parameters in the parametrization of the expected photon density have been updated. Most importantly, it allows for larger upward fluctuations by folding the likelihood description (Equation 4.5) with a skewed Student's t-distribution in $\ln(x)$ [135]

$$G_\Lambda(x) = \frac{const.}{x} \cdot (e^{-\omega y} + (y/\sigma)^2)^{-1}\,, \tag{4.6}$$

$$\text{with} \quad y = \ln(x/\Lambda)\,, \tag{4.7}$$

with the *skewness* parameter ω. This covers high-energy stochastics along the muon track and leads to an improved energy resolution as shown in Figure 4.1.

4.1.1.3 Truncated Energy

Truncated Energy (TE) applies the Monte Carlo approach introduced previously, but with a modification to handle large stochastic energy losses. It divides the muon trajectory into up to 15 bins, split by planes perpendicular to the track (*Bin mode*). An alternative mode defines one bin for every DOM (*DOM mode*) [81]. For each bin, dE_μ/dx is evaluated and the bins with the largest values are omitted. A cut percentage of 40 % for the Bins method and 50 % for the DOMs mode have been found to be the ideal values [81]. Afterwards, dE_μ/dx is evaluated for the reduced set of DOMs which gives the improved energy resolution displayed in Figure 4.1. A modification also considers regions/DOMs without photon hits and is named *AllBins/AllDOMs*[1].

4.1.2 Angular resolution estimation

The estimation of the angular reconstruction accuracy of individual events is twofold important in experiments like IceCube. First, it can be used to identify misreconstructed tracks and improve the background rejection. Second, in unbinned maximum likelihood point source searches it serves as standard deviation in the Gaussian spatial PDF as specified in Section 5.2.

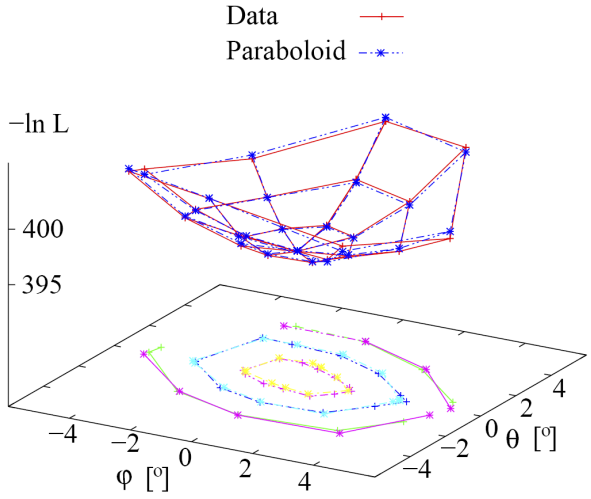


Figure 4.3: The paraboloid in blue has been fit to the reference points (red) spread in the zenith θ and azimuth ϕ plane around the optimum. The z axis shows the negative logarithm of the likelihood [137].

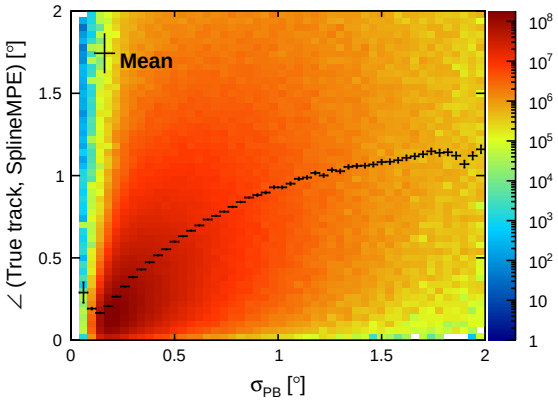


Figure 4.4: The Paraboloid fit error estimation σ_{PB} and the space angle between true Monte Carlo track and SplineMPE reconstruction are correlated with a correlation factor of 0.56. The mean value in every σ_{PB}-bin is shown black.

High level reconstructions (see Section 3.1) find the most probable track by converging towards a minimum in the negative logarithmic likelihood landscape. The accuracy of the

[1]The modifications give a slightly improved resolution at lower energies [136].

resulting trajectory can be evaluated by exploring the shape of that landscape around the optimum. A narrow, well defined minimum indicates a good resolution while a broad, sketchy surrounding would point to a bad reconstruction.

In the ideal Gaussian case, the minimum can be described by a paraboloid. This paraboloid is fit to a set of minimized reference points around the given optimum with respect to the track angles zenith Θ and azimuth Φ [137]. This process is illustrated in Figure 4.3. The accuracy of the track, also named Paraboloid sigma, can be evaluated afterward from the semi-axis e_1 and e_2 of the confidence ellipse of the paraboloid:

$$\sigma_{\mathrm{PB}} = \sqrt{\frac{e_1^2 + e_2^2}{2}}\,. \tag{4.8}$$

Figure 4.4 confirms the correlation of the Paraboloid fit estimation with the space angle between reconstructed and true Monte Carlo track.

4.1.3 IceTop veto

The wide spread air showers creating IceCube's main background of atmospheric muons frequently create hits within the IceTop surface array. Assuming a shower front perpendicular to the in-ice track reconstruction, the number of IceTop hits within a time window around the shower transit can be evaluated. This number of IceTop veto hits N_{Veto}, coincident with the in-ice hits, can be used for very efficient background rejection.

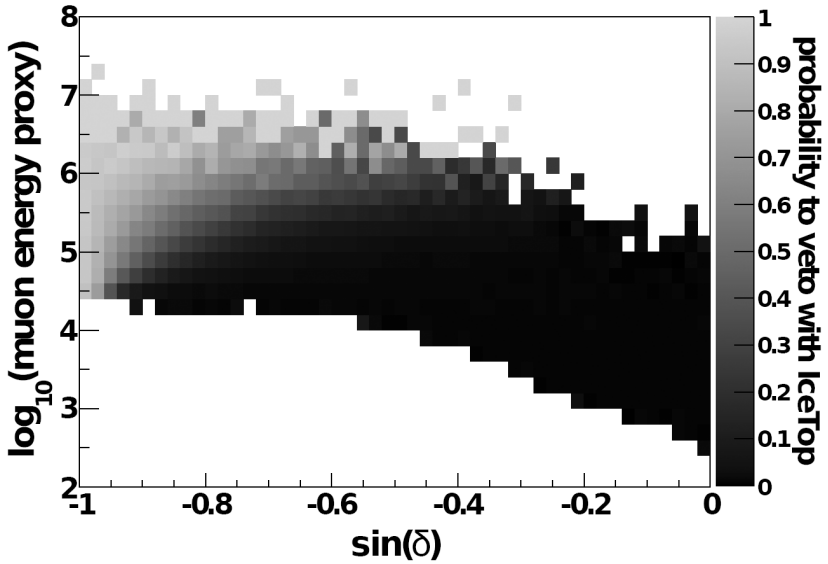


Figure 4.5: The IceTop veto shown here rejects atmospheric events with more than three coincident veto hits N_{Veto} and is most efficient at high energies and vertical directions, reaching a background reduction of over 90 % [109]. δ denotes the declination.

As shown in Figure 4.5, the veto, here rejecting events with $N_{\mathrm{Veto}} \geq 3$, is most efficient at high energies and vertical directions sorting out over 90 % of background. The probability

for accidental coincidences has been examined from off-time window data rates to be below 1 % for all energies and declinations [109]. To further decrease the accidental coincidence rate, the veto can be restricted to the most efficient regions in the declination-energy parameter space. This is done for the IC79b sample in Section 4.3.4.

4.1.4 Hit topology variables

Number of channels and strings: The number of channels N_{Ch} is the number of DOMs that registered a hit after hit cleaning and N_{Str} gives the number of strings with at least one hit DOM after a possible hit cleaning. Both are closely connected to the light deposit in the detector and the lepton energy.

Depth of hits: Z_{travel} is the difference between the average depth of the first quartile of hits in time and the average depth of all hits. It therefore measures the vertical extent and orientation of the track. Z_{sigma} denotes the standard deviation of the z coordinates of all channels and can be used to normalize Z_{travel} for a better comparison.

Charge values: The largest measured DOM charge in an event is called Q_{max} and the sum of all DOM charges in an event is called Q_{tot}. The ratio of the two separates tracks and cascade like events, as tracks create a more distributed light pattern.

Center of gravity (CoG): The DOM charge weighted center of gravity in the z coordinate Z_{CoG} can help to identify atmospheric muons as they are more likely to be found in the upper part of the detector. The CoG radius is a containment variable defined with the center of gravity in x and y coordinates as $R_{\mathrm{CoG}} = \sqrt{X_{\mathrm{CoG}}^2 + Y_{\mathrm{CoG}}^2}$. Contained events usually have a better reconstruction accuracy.

4.1.5 Hit-track topology variables

Hit distribution along the track: The largest distance along the reconstructed track without hits within a cylinder with a radius of 150 m is called L_{Empty}. This value can be large for two coincident muons falsely reconstructed as one track. The distance between the center of gravity of the first quartile of hits in time and the last quartile of hits in time along the reconstructed track is called separation L_{Sep}. This value should be large for high-energy muons.

Direct hits: A photon hit is considered as direct if the time residual lays between -15 ns and 75 ns. N_{Dir} is the number of direct hits for the reconstructed track while L_{Dir} is the maximum of all distances between two direct hits. Both values are typically large for well reconstructed high-energy muons.

Reconstructed track distances: The distance between the DOM charge weighted center of gravity of the hits used for reconstruction and the reconstructed track itself is called L_{CoG} and can be large for coincident events and tracks clipping the edges of the detector. For high-energy muons nicely traveling through the detector this value is usually small. The closest approach to the center of the detector L_{CoD} is the distance between the reconstruction and the origin (0,0,0) of the IceCube coordinate system in the middle of the detector. The smaller the value the more contained the event is.

4.1.6 Reconstruction quality variables

Reduced likelihood: The negative reduced logarithmic likelihood value found as optimum in a reconstruction is a quality parameter for the fit as it is small for well reconstructed tracks and large for misreconstructed events. Following the reduced χ^2 approach, it is defined from the optimal likelihood result $\mathcal{L}$ as $\mathcal{L}_{\mathrm{reduced}} = -\log(\mathcal{L})/(N_{\mathrm{Ch}} - N_{\mathrm{dof}})$. For historic reasons, with two track angles and three track vertex coordinates, the number of degrees of freedom N_{dof} is five. However, it was found that other variants with e.g. $N_{\mathrm{dof}} = 3.5$ and $N_{\mathrm{dof}} = 2.5$ give a better discrimination power for soft spectra and are less energy dependent. These variants are then named $\mathcal{L}^{(3.5)}_{\mathrm{reduced}}$ and $\mathcal{L}^{(2.5)}_{\mathrm{reduced}}$.

Line-fit velocity and smoothness: The line-fit velocity V_{LF} defined in Section 3.1.1 is sensitive on the light distribution signature, and can identify long muon tracks. The smoothness parameter S_{all} measures how equal all photon hits are distributed along the track.

4.1.7 Reconstruction comparisons

High-noise reconstruction: Usually, likelihood reconstructions use a flat noise probability of $10^{-8}\,\mathrm{ns}^{-1}$ as discussed in Section 3.5.3. Raising this value by a factor of 10^5 can swamp disadvantageous noise hits and even coincident tracks and turn the reconstruction result into a largely different direction. Although the overall resolution of this configuration is much worse, the angular difference $\Psi_{\text{high-noise}}$ to the original reconstruction is a powerful tool to detect misreconstructed and coincident events.

Split reconstructions: Splitting the hits of an ideal high-energy muon into two parts by time and by geometry should result in four subsets of hits that reconstruct to roughly the same direction as the whole event. For badly reconstructed or coincident events this is commonly not the case. For this value, the four subsets of an event are reconstructed and the angular distance Ψ_{split} between the smallest (the most downgoing) zenith angle and the zenith angle of the whole event's reconstruction is evaluated. A big value can identify coincident, downgoing muons.

4.1.8 Status flags

Some of the methods presented here might fail in rare cases due to different reasons. These could be missed preconditions or a failed likelihood optimization. If the result of a method is essential for the continuative process, failed events have to be cut. The effects of the particular status cuts are described later where they are applied.

4.2 Basic data processing

IceCube's data acquisition, described in Section 2.5.3, creates a data rate of 1 TB per day. It has to be reduced at South Pole to fit the limited satellite bandwidth of 100 GB per day [138]. Additionally, CPU intensive methods cannot be applied on such vast event rates. Thus the data rate has to be reduced in several processing and selection levels (L), where the complexity of the methods rises with decreasing rates. Trigger, calibration and hit extraction, covered in Section 2.5.3, are commonly seen as level 0 (L0). The further data processing in IceCube combines numerous methods and applications on various data streams designed for different analyses and physics goals. To avoid confusion, the subsequent processing stages described below, solely focus on the steps and methods relevant for this work.

4.2.1 L1 and L2: filtering and first reconstructions

The L1 and the L2 processing are combined into one processing effort. The input from L0 provides triggered events[2] composed by a series of photon hits with a timestamp and a charge. On L1, a time window cleaning is applied on the hits that passed the HLC conditions, only keeping hits within a span of 6 µs around the trigger time. This further reduces the number of noise pulses. A line-fit is performed on these time window cleaned pulses serving as a starting track for a PandelSPE reconstruction. Afterwards, the data rate is reduced by a muon filter requiring PandelSPE $\mathcal{L}^{(2)}_{\text{reduced}} < 8.1$ for upgoing reconstructed tracks with $\cos(\Theta) < 0.2$, where Θ is the PandelSPE zenith angle. This strongly suppresses falsely reconstructed atmospheric muons. In the downgoing region, a cut on the total event charge Q_{tot} discards low energetic atmospheric muons depending on the reconstructed zenith angle

$$\log(Q_{\text{tot}}) > u \cdot \cos\Theta + v\,, \tag{4.9}$$

with $u = 3.9$ and $v = 0.55$ for $0.2 < \cos\Theta < 0.5$, and $u = 0.6$ and $v = 2.2$ for $\cos\Theta >= 0.5$. The data passes the muon filter at a rate of ≈ 40 Hz.

Afterwards, on L2, a more sophisticated hit cleaning method is applied by adding hits that missed the HLC conditions (SLC hits). Even though the HLC cleaning very efficiently discards PMT noise hits, it also removes about 30 % of the track related pulses [139]. In

[2] Containing a data rate of ≈ 2.2 kHz from the simple multiplicity trigger, relevant for this work.

the seeded RT cleaning, all HLC hits are looped and a SLC hit is added, if it appears within a certain radius R and a certain time window T of the center of gravity of the HLC pulses. If no SLC hit complies with the condition, the requirement is expanded and an SLC hit is added if it is within the RT range of any HLC hit. The procedure can be iterated until no more hits are added. Here, $R = 150\,\mathrm{m}$, $T = 1000\,\mathrm{ns}$ and two iterations are chosen, keeping 95 % of the signal related hits while only slightly increasing the noise rate [139]. Finally, the same time window cleaning as on L1 is used and the result, named TWSRTPulses, is the base for all reconstructions described below this point.

During the L2 processing, a PandelMPE reconstruction, seeded with a previous PandelSPE fit, is performed. The energy of the resulting track is then estimated by the MuE method. Finally, the output is transmitted to the northern hemisphere via satellite.

4.2.2 L3: advanced routines and cuts

The processing on level 3 directly aims for high-energy point source searches and thus contains specifically tailored cuts for this purpose. Additionally, all values needed in point source analyses are evaluated here. As a result, the output of L3 is also the starting point of the reprocessing of the IC79 data, which is part of this work and adds the SplineMPE reconstruction described in Chapter 3.

In a first step, a set of cuts based on the PandelMPE reconstruction is applied on all data that passed the muon filter

$$\mathrm{DI} > 2.0 \;\;\&\;\; \mathcal{L}^{(3.5)}_{\mathrm{reduced}} <= 12.0 \;\;\&\;\; L_{\mathrm{Sep}} > 0.0 \;\;\&\;\; L_{\mathrm{Empty}} <= 500.0 \tag{4.10}$$

$$\text{with} \quad \mathrm{DI} = (L_{\mathrm{Dir}}/60.0)^2 + (N_{\mathrm{Dir}}/15.0)^2\,. \tag{4.11}$$

In words, the cuts require that an event has a descent number of direct hits, continuously distributed along a long trajectory with a proper likelihood reconstruction. For the southern hemisphere, another more sophisticated energy cut is implemented to reduce low energetic atmospheric muons. It depends on the PandelMPE zenith angle Θ and the total event charge Q_{tot}

$$\log(Q_{\mathrm{tot}}) >= 1.05 \cdot (0.76 + 4.88 \cdot \cos(\Theta) - 4.78 \cdot \cos(\Theta)^3 + 2.13 \cdot \cos(\Theta)^5)\,. \tag{4.12}$$

Both cuts result in a data rate of $\approx 2.2\,\mathrm{Hz}$ in the downgoing branch and $\approx 0.9\,\mathrm{Hz}$ for upgoing muons. A signal efficiency between 80 % and 90 % is achieved, depending on the spectrum.

Afterwards, a topological trigger (TTrigger) tries to split coincident tracks depending on the hit distribution in time and space. If an event was split, the MPE and MuE routines from L2 are repeated on the sub tracks. If the sub events pass the L3 cuts above, they are kept and discarded otherwise.

The reconstructions relevant for this work which are performed on L3 on all events and potential sub events, are listed in Table 4.1. In addition to the PandelMPE performed on L2, the same fit is repeated with a noise probability raised by a factor of 10^5 as motivated

in Section 4.1.7. The accuracy of both MPE fits is evaluated with the Paraboloid method. Based on the PandelMPE track, the energy estimators MuEX and TE are performed. For upgoing tracks, additional routines are used. The hit pattern is split evenly in time and geometry and the resulting sub events are fit with a 4 iterations PandelSPE (see Section 4.1.7).

All events	Upgoing events
PandelMPE Paraboloid	Geometry split PandelSPE, 4 iterations
High-noise PandelMPE	Time split PandelSPE, 4 iterations
High-noise PandelMPE Paraboloid	
PandelMPE MuEX	
PandelMPE Truncated Energy (TE)	

Table 4.1: Reconstructions relevant for this work performed on L3 for all events and upgoing events only.

4.2.3 L3b: Adding the SplineMPE reconstruction

The reprocessing[3] of the IC79 data, which has been performed as part of this work, uses the result of L3 from Section 4.2.2 as a starting point. The ultimate goal is a final neutrino sample featuring the advantages of the SplineMPE reconstruction investigated in Section 3.4. Therefore, along with the SplineMPE fit, all values introduced in Section 4.1 have to be reevaluated based on the new reconstruction. From here on, the creation of the SplineMPE analysis sample IC79b is specified and the original IC79 set from [109] is used as comparison.

The SplineMPE itself is performed in the standard setting with a flat noise level of 10 Hz. The improvements developed in Chapter 3.5 could not be activated due to the lack of computational resources as they significantly slow down the reconstruction. The slowest component in the chain is the Paraboloid fit because it has to optimize the likelihood at several reference points around the optimum. For experimental data and all Monte Carlo data sets, the reprocessing took several weeks on 300 CPU cores (Intel Xeon E5345 @ 2.33GHz).
All reconstructions in the process are based on the time window cleaned and seeded RT cleaned DOM hits (TWSRTPulses). The first step in the data flow, displayed in Figure 4.6, is a PandelMPE reconstruction based on the bootstrapping technique described in Section 3.1.3. Below 10 TeV it delivers a better angular resolution than PandelMPE which, however, performs better at higher energies. As a result, both fits together serve as a good seed for the SplineMPE evaluation, which is therefore performed twice with different starting parameters storing only the optimal choice.

Based on the result of the new reconstruction, the three energy estimators MuE, MuEX and TE are applied as well as the angular uncertainty routine Paraboloid. Additional

[3]The software setup is specified in Table B.3 of the appendix.

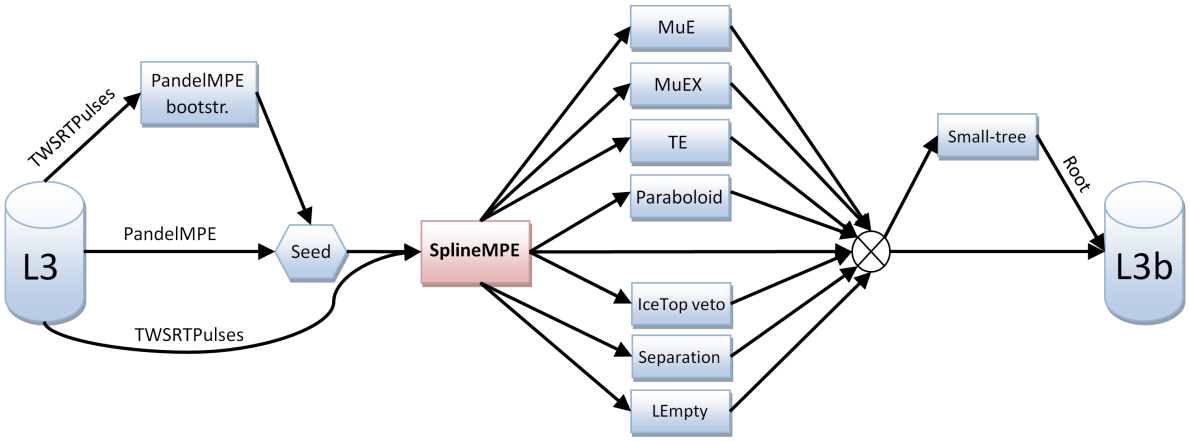


Figure 4.6: L3b data flow: The SplineMPE reconstruction is seeded with one PandelMPE with bootstrapping and the L3 PandelMPE without bootstrapping. Afterwards, all track related variables are calculated and the result is stored as ROOT file and in the IceCube file format.

modules calculate L_{Sep}, L_{Empty} and the IceTop veto (Section 4.1.3). Finally, the whole data stream is passed to a module[4] which not only calculates the missing values specified in Section 4.1, but also writes output files in the format of CERN's Data Analysis Framework Root [123]. Simultaneously, the output is also saved in the IceCube file format. At this point in L3b, the data are still dominated by atmospheric muons. The reduction to atmospheric neutrino level in the upgoing region and to high-energy atmospheric muon level in the downgoing hemisphere, is performed with boosted decision trees and described in Section 4.3.

4.3 Multivariate classification with boosted decision trees

Classification of signal events from high statistics data sets is an ideal task for multivariate machine learning methods such as the well-established artificial neural networks (ANN) or boosted decision trees (BDT). These concepts usually perform superior compared to sets of manually adjusted cuts and result in an improved efficiency. In this thesis, BDTs have been chosen for the definition of the final neutrino sample, similarly to what was done in the original IC79 analysis [109]. While the theoretical performance of ANNs is expected to be better in many cases in principle, BDTs regularly outperform these in practical use. The reason is the robustness of the BDT classification with respect to chosen structure, input variables and number of training samples [140].

BDT training: Usually the available data is split into a training sample and an independent test sample for performance evaluation, both consisting of signal (S, astrophysical

[4] named Small-tree

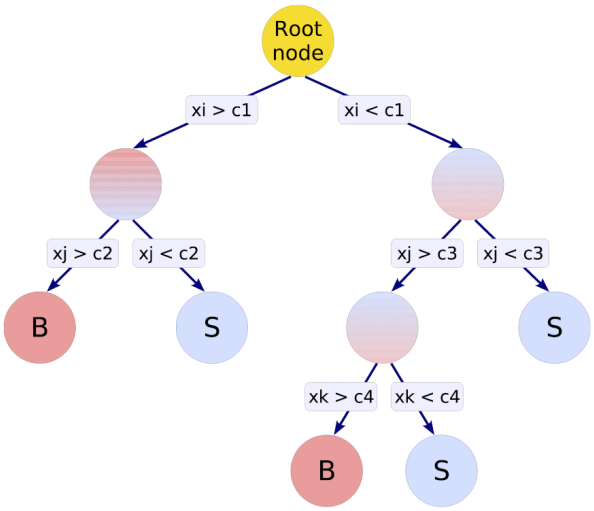


Figure 4.7: The BDT structure is formed by splitting the training sample recursively by the best possible cut c of an input variable x until a stopping condition is met. The bottom nodes (leaves) are then labeled as signal (S) or background (B), depending on the majority of events in that respective node [140].

neutrinos) and background (B, atmospheric particles) events. A tree is formed in a recursive process starting with the whole training sample. All input variables are scanned for the best discrimination between signal and background and the set is split into two sub nodes by the best possible cut at this point[5]. For the two created sub containers this process is repeated until a stopping condition is reached. This can be the maximal depth of the tree or the minimal number of events in a node which are the main settings for controlling overtraining[6]. The resulting, lowest nodes (leaves) are then classified as signal or background depending on which class of events is the majority [140]. The tree structure is illustrated in Figure 4.7. Once the tree is fixed an event defined by a tuple of input parameters can be classified by descending through the tree structure into a signal or background leaf.

Boosting: A single decision tree is a rather weak classifier and prone to statistical fluctuations. These can change the whole tree structure if a different cut variable is chosen close to the root due to slight changes in the training events. A process called Adaptive Boosting increases the stability and performance of the classifier.

After the creation of the first tree, higher weights are assigned to events that ended up in the wrong leaf – e.g. a signal event in a background dominated leaf – and a second tree is build with the reweighted training sample. This process is usually repeated a few

[5]The best cut is identified by the Gini index defined as $p \cdot (1-p)$ with the purity $p = S/(S+B)$ [140].

[6]Overtraining occurs if the classifier structure is too complex for the given training data. As a result, the BDT starts to "memorize" single events instead of common properties of a class which leads to an excellent performance in the training sample and poor results on new, unknown data. Overtraining is monitored by executing the classifier on the independent test sample.

hundred times forming a forest of boosted decision trees. A new event weight is obtained by multiplying the misclassified event by a boost weight α derived from the misclassification rate, R, of the previous tree. The training rate can be adjusted by the exponent β [140]:

$$\alpha = \left(\frac{1 - \mathrm{R}}{\mathrm{R}}\right)^{\beta} \quad . \tag{4.13}$$

The whole sample is then normalized again to keep the sum of weights constant. The result $y_{\mathrm{BDT}}(x)$ of such a forest of decision trees for a set of input variables x is calculated by [140]

$$y_{\mathrm{BDT}}(x) = \frac{1}{N_{\mathrm{trees}}} \cdot \sum_{i}^{N_{\mathrm{trees}}} \ln(\alpha_i) \cdot h_i(x)\,, \tag{4.14}$$

where N_{trees} is the number of trees in the forest and $h_i(x)$ is the response from one single tree. As this basic tree result $h_i(x)$ is defined as $+1$ for signal and -1 for background, the overall result of the forest $y_{\mathrm{BDT}}(x)$ is large for signal and low for background events.

Pruning: *Cost complexity pruning* is one way to reduce possible overtraining of the tree by cutting insignificant leaves. The decrease of the misclassification rate R in the subtree below a node compared to the node itself is related to the number of nodes in the subtree [140]:

$$\rho = \frac{R(\text{node}) - R(\text{subtree below that node})}{N_{\mathrm{nodes}}(\text{subtree below that node}) - 1}\,. \tag{4.15}$$

The node with the lowest cost complexity ρ is removed recursively, while ρ is smaller than an adjustable *PruneStrength*.

4.3.1 BDT setup

The multivariate classification in this work is using the TMVA package [140], which is implemented in the ROOT Analysis Framework [123]. To preserve the possibility for direct comparison, the basic BDT setup from the original IC79 sample [109] was used. The sky is divided into 3 different zenith angle regions defined by the SplineMPE reconstruction. In each of them, different requirements for background rejection are imposed:

- $0°$ **to** $90°$**:** In this downgoing region, the main task is the separation of atmospheric muons and astrophysical neutrinos. This is mainly done by a zenith dependent energy cut and an IceTop veto. Additionally, poorly reconstructed tracks have to be identified and discarded.

- 85° **to** 130°**:** The horizontal upgoing range has to be cleaned from slightly false reconstructed atmospheric muons as well as from falsely reconstructed coincident tracks. The overlap with the downgoing zone of 5° is chosen for two reasons. First, the region where atmospheric muons are shielded actually extends up to a zenith angle of 85°, as IceCube is located 2000 m below the surface. Second, because these two zenith regions are very different, the overlap is needed for a smooth transition of the output data. Only well reconstructed non-coincident tracks shall be retained.
- 130° **to** 180°**:** In this zone, the background is already very low and mainly consists of falsely fit coincident events and a few heavily misreconstructed atmospheric muons. Astrophysical signal events are suppressed at energies above 1 PeV where the Earth is becoming opaque for neutrinos (see Figure 4.13).

Because of the atmospheric muon background, only at higher energies IceCube is sensitive to hard neutrino spectra in the southern hemisphere. However, in the upgoing regions the event selection should consider softer spectra in addition. Consequently, in these two zenith ranges, one BDT is trained with an E^{-2} neutrino spectrum as signal and a second BDT with a softer spectrum including an exponential cutoff: $E^{-2.4} \cdot \exp(-E/7\,\mathrm{TeV})$.

Region	Zenith range	Soft spectrum signal	Hard spectrum signal
Downgoing	0° to 90°	-	BDT1, BDT2
Horizontally upgoing	85° to 130°	BDT1, BDT2	BDT1, BDT2
Vertically upgoing	130° to 180°	BDT1, BDT2	BDT1, BDT2

Table 4.2: The BDT setup for the IC79b final neutrino sample consists of 10 BDTs in 3 zenith regions. Besides the hard E^{-2} spectrum in all orientations, the upgoing zones also train a softer cutoff spectrum, parametrized by $E^{-2.4} \cdot \exp(-E/7\,\mathrm{TeV})$. For training stability reasons, each BDT is split into a BDT1 and a BDT2 (see below).

A reweighted Monte Carlo muon neutrino event set[7] is used to model the signal for the BDT training. It is was produced with the PPC photon propagator and the Spice Mie ice model (Section 2.5.4). For the background, the actual experimental data from the L3b data set can be used, as the expected signal ratio is still negligibly small. The possibility to use real background avoids potential systematic uncertainties and weaknesses introduces by using simulation. For blindness reasons only 10 % of the whole year are used in the BDT training. This restriction reduces the available statistic and can lead to slightly unstable training results when using many input variables (degrees of freedom). Hence, every BDT is split into a BDT1 and a BDT2 trained and evaluated separately with different input variables. The result is then combined by:

$$\mathrm{BDT} = (\mathrm{BDT1} + 1) \cdot (\mathrm{BDT2} + 1)\,. \tag{4.16}$$

The whole setup contains a total of 10 BDTs and is summarized in Table 4.2.

[7] Internal IceCube data set number 6359.

4.3.2 Input variables

The choice of BDT input variables is a process with some trial and error component. A general approach first starts with the selection of variables providing the best discrimination between signal and background. This can be done by evaluating the separation $\langle S^2 \rangle$ of all variables given by [140]

$$\left\langle S^2 \right\rangle = \frac{1}{2} \int \frac{(\widehat{y}_S(y) - \widehat{y}_B(y))^2}{\widehat{y}_S(y) + \widehat{y}_B(y)} dy \,, \tag{4.17}$$

where $\widehat{y}_S(y)$ and $\widehat{y}_B(y)$ are the signal and background PDFs. A separation of zero reflects identical signals while a value of one indicates perfect division. The next step is the reduction of the number of variables by discarding variables with strong correlations. This part goes along with steady BDT training and evaluation on the test sample to monitor the resulting efficiency.

For this work, the variables chosen in the original IC79 selection were replaced by their counterparts based on the SplineMPE reconstruction. This proceeding ensures the direct comparability of both final samples with respect to the impact of the new reconstruction.

BDT 1	BDT 2
PandelMPE $\Psi_{\text{high-noise}}$	Z_{CoG}
PandelMPE $\mathcal{L}^{(3.5)}_{\text{reduced}}$	PandelMPE Ψ_{split}
V_{LF}	$Z_{\text{travel}}/Z_{\text{sigma}}$
R_{CoG}	$-\log(Q_{\text{max}}/Q_{\text{tot}})$
SplineMPE $\mathcal{L}^{(2.5)}_{\text{reduced}}$	SplineMPE L_{Sep} / SplineMPE L_{Dir}
SplineMPE L_{CoG}	SplineMPE σ_{PB}
log(SplineMPE MuEX)	SplineMPE N_{Dir}
abs(S_{all})	SplineMPE L_{CoD}
SplineMPE L_{Empty} / SplineMPE L_{Dir}	

Table 4.3: Input variables of the upgoing BDTs.

However, a few exceptions were made. For the upgoing BDTs, the SplineMPE zenith angle had to be removed as it created a sharp rate drop below the horizon by cutting low energy tracks. While this is the expected behavior to reduce atmospheric muons leaking in, it was not possible afterward to assure a smooth transition to the downgoing region. Additionally, the use of MuEX instead of TE in the upgoing hemisphere gave a slightly improved efficiency. Furthermore, the $\mathcal{L}^{(3.5)}_{\text{reduced}}$ value based on the Pandel function is kept because it gives a slightly better background rejection for very hard cuts compared to SplineMPE $\mathcal{L}^{(2.5)}_{\text{reduced}}$. The final choice of input variables is listed in Table 4.3 and Table 4.4 for the upgoing, respectively downgoing hemisphere.

BDT 1	BDT 2
PandelMPE $\Psi_{\text{high-noise}}$	Z_{CoG}
PandelMPE $\mathcal{L}^{(3.5)}_{\text{reduced}}$	$Z_{\text{travel}}/Z_{\text{sigma}}$
V_{LF}	$-\log(Q_{\text{max}}/Q_{\text{tot}})$
R_{CoG}	SplineMPE L_{Sep}
SplineMPE $\mathcal{L}^{(2.5)}_{\text{reduced}}$	SplineMPE σ_{PB}
SplineMPE L_{CoG}	SplineMPE N_{Dir}
log(SplineMPE TE)	SplineMPE L_{CoD}
abs(S_{all})	
SplineMPE zenith	

Table 4.4: Input variables of the downgoing BDTs.

Figure B.5 of the appendix shows examples for the signal-background separation achieved by single variables before they are combined into a strong classifier by the BDTs.

4.3.3 BDT configuration

As stated in Section 4.3, the BDT performance is relatively robust w.r.t. changes in the tree structure. Nevertheless, small improvements could be possible, so different settings were tested alongside with the BDT configuration of the original IC79 sample. Table 4.5 shows the weighted signal efficiencies of different BDT settings at three background efficiencies for the downgoing BDT1[8]. The BDTs of the older IC79 sample are identifiable by the activated pruning, which is not recommended any more [140]. To avoid overtraining, the configurations without pruning are set up with a minimum number of events within a node of 400 instead of 100.
From these results it is obvious that an improvement is possible compared to the original IC79 BDT configuration. Therefore, $N_{\text{trees}} = 800$ and depth $= 4$ was chosen for the downgoing BDTs, because higher values for the maximal depth increased the effect of overtraining. The minimum number of events is set to 400 and adaptive boosting with a learning rate of $\beta = 0.5$ is used. In the upgoing hemisphere, the same tests showed no improvement in the background rejection so the BDT configuration of the old IC79 sample is used. Table 4.6 lists the final BDT settings.

[8]The efficiencies for BDT2 can be found in Table B.2 of the appendix.

N_{trees}	Maximal depth	Prune strength	Signal efficiency at bkg eff.(error): @B=0.01	@B=0.10	@B=0.30
800	5	-	0.605(03)	0.822(02)	0.931(01)
800	4	-	0.598(03)	0.821(02)	0.931(01)
400	5	-	0.600(03)	0.821(02)	0.930(01)
400	6	-	0.599(03)	0.810(02)	0.927(01)
800	6	-	0.601(03)	0.811(02)	0.927(01)
800	3	-	0.592(03)	0.817(02)	0.928(01)
400	4	-	0.594(03)	0.815(02)	0.928(01)
400	3	-	0.587(03)	0.816(02)	0.926(01)
800	3	5	0.564(03)	0.806(02)	0.920(01)
800	2	-	0.565(03)	0.801(02)	0.919(01)
400	3	5	0.563(03)	0.799(02)	0.914(01)
400	2	-	0.559(03)	0.798(02)	0.912(02)

Table 4.5: Signal efficiency of different BDT configurations for the downgoing BDT1 at three different background efficiencies evaluated for the independent test sample. The BDTs from the original IC79 sample are identifiable by the activated pruning.

Hemisphere	N_{trees}	Max. depth	Prune method	Prune strength	Boost method	Boost rate β	Min. events per node
downgoing	800	4	-	-	adaptive	0.5	400
upgoing	800	3	Cost Compl.	5	adaptive	0.5	100

Table 4.6: BDT configuration used in this work.

4.3.4 Pre-classification cuts

Before the L3b data is fed to the BDTs, a set of pre-cuts discards events that are obviously useless or failed on important reconstructions. Again, the original IC79 variables were replaced by their SplineMPE counterparts. The following pre-cut was chosen:

$$\begin{aligned} &\text{SplineMPE Paraboloid status} = \text{OK} \;\;\&\;\; \text{SplineMPE TE status} = \text{OK} \\ \&\; &\text{SplineMPE } \sigma_{\text{PB}} <= 6.0 \;\;\&\;\; \text{SplineMPE } N_{\text{Dir}} >= 3.0 \\ \&\; &N_{\text{Ch}} >= 10 \;\;\&\;\; N_{\text{Str}} >= 2.0\,. \end{aligned} \tag{4.18}$$

The result of the Paraboloid fit is an important part of the likelihood analysis method introduced in Section 5.2, thus the status of the fit needs to be ok[9]. For the same reason, the TE status is checked. TE fails if there are less than three 80 m bins with hits in a 150 m cylinder along the track. Hence, as a side effect, this status flag selects track like events and discards cascade like events from NC interactions. The other cuts discard remaining events with too few hits, strings or direct hits as well as an angular error estimation above 6° which would give a negligible event weight in a point source analysis.

For the Monte Carlo events used as training signal, two additional cuts are applied. Neutrinos with an energy below 200 GeV are discarded because the neutrino-muon interaction angle reaches the reconstruction accuracy (Equation 2.16). Additionally, as the BDTs should not be trained with badly reconstructed events, tracks with a reconstruction error larger than 5° are cut.

Finally, the IceTop veto introduced in Section 4.1.3 is applied on the data, discarding events depending on estimated energy (TE), reconstructed zenith angle and IceTop data. In this case, the configuration of the original IC79 sample remains unchanged as small improvements in the angular reconstruction are negligible for the veto. The veto is defined as

$$\begin{aligned} &\neg(\cos(\Theta) > 0 \quad \&\;\; \log(\text{TE}) > 5 \quad \&\;\; N_{\text{Veto}} >= 3) \\ \&\; &\neg(\cos(\Theta) > 0.8 \quad \&\;\; \log(\text{TE}) > 4 \quad \&\;\; N_{\text{Veto}} >= 3) \\ \&\; &\neg(\cos(\Theta) > 0.9 \quad \&\;\; N_{\text{Veto}} >= 3)\,, \end{aligned} \tag{4.19}$$

where Θ is the zenith angle of the PandelMPE reconstruction and N_{Veto} is the number of IceTop veto hits.

The upgoing (downgoing) cut efficiency of these pre-cuts is 9.7 % (68.3 %) for the data and 74.5 % (77.4 %) for an E^{-2} signal. Table 4.7 shows the data rates after the cuts and the changes compared to the original cuts based on the Pandel function for the whole sky and the three subregions[10]. While the signal efficiency is only changed slightly, the data

[9] The failure rate for the Paraboloid fit is between 1 % and 2 % and mostly caused by badly reconstructed tracks.

[10] The sum of the rates in the subregions do not match the all sky rate because of the overlap between 85° to 90° zenith angle.

Zenith region	Data rate	Change w.r.t old IC79 sample E^{-2} signal	Data
All Sky	1.85 Hz	−0.7 %	−8.9 %
Downgoing	1.79 Hz	−1.8 %	−7.0 %
Horizontal upgoing	0.07 Hz	−0.8 %	−37.7 %
Vertical upgoing	0.01 Hz	+0.7 %	−60.7 %

Table 4.7: Data rates of the pre-cuts in the three zenith angle regions and for the whole sky. Rate changes with respect to the old IC79 are given for signal and background.

rates have been decreased up to 60 % in the vertical upgoing zone compared to the PandelMPE based sample. The reason is that the SplineMPE reconstruction, seeded with the bootstrapping PandelMPE, identifies and corrects many previously upgoing reconstructed atmospheric muons to be downgoing.

The BDTs were trained with 10 % of the pre-cut data (background) and Monte Carlo (signal). Section 4.3.5 describes the strategy of finding the best cut values for the BDT response values and how to combine them into a smooth all sky sample.

4.3.5 BDT cut optimization

After BDT training, the optimal cut value on the BDT response derived with Equation 4.16 has to be evaluated. Figure 4.8 shows this BDT output for different data types for the trees trained with the hard spectrum in the horizontal region. Most of the data in black and the simulated atmospheric muons in blue are located at smaller values, indicating background. The disagreement in rate is caused by a systematic underestimation of coincident events in the IceCube simulations. At output values of 1.2, the simulated atmospheric neutrinos in bright blue start to very well match the data while the muon background decreases quickly. The expectations for example neutrino signal fluxes of $2 \times 10^{-7} \cdot (E/\mathrm{GeV})^{-2}$ and $1 \times 10^{-12} \cdot (E/\mathrm{GeV})^{-1}$ (both in $\mathrm{GeV}^{-1}\,\mathrm{m}^{-2}\,\mathrm{s}^{-1}$) are shown in orange and red.

The final sample is optimized for a discovery potential on a 5σ level as defined in Section 5.2.1 for soft and hard spectra. The optimization is done separately for every zenith region and has to comply with the following conditions:

- optimal discovery potential for hard and soft spectra,
- both sub BDTs give signal-like response: $BDT1 > 0 \;\&\; BDT2 > 0$,
- smooth event rate and energy distribution transitions at zenith range edges[11].

In a first step, the discovery potential for an E^{-2} spectrum is evaluated for different BDT cuts (see Equation 4.16) at a chosen zenith angle for the BDTs trained with the E^{-2}

[11] Hard edges at the BDT transitions in energy distribution or event rate can cause artifacts in maximum likelihood ratio analyses as introduced in Section 5.2 and have to be avoided.

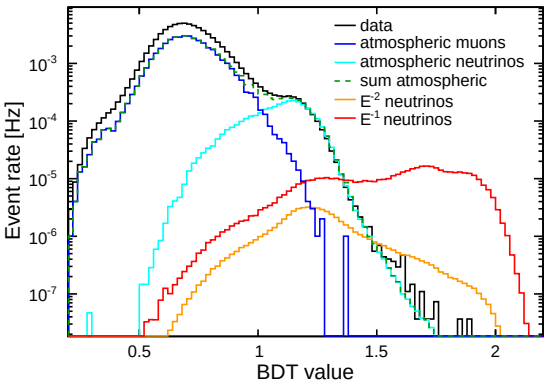


Figure 4.8: The plot shows the results from the hard spectrum BDT in the horizontal region for data (black), atmospheric muons (blue), atmospheric neutrinos (cyan), and two neutrino fluxes of $2 \times 10^{-7} \cdot (E/\mathrm{GeV})^{-2}$ (orange) and $1 \times 10^{-12} \cdot (E/\mathrm{GeV})^{-1}$ (red) (both in $\mathrm{GeV}^{-1}\,\mathrm{m}^{-2}\,\mathrm{s}^{-1}$).

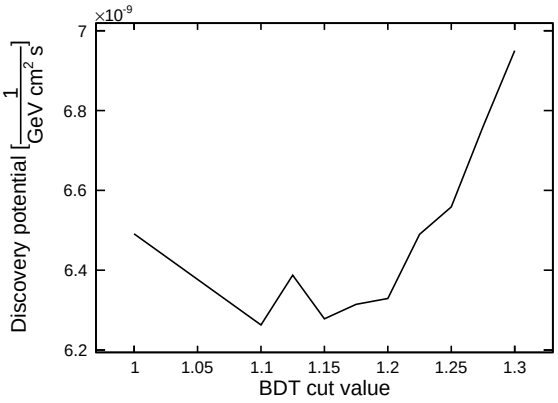


Figure 4.9: The plot shows the discovery potential given as scaling factor for an $(E/\mathrm{GeV})^{-2}$ spectrum for different cut values of the horizontal upgoing hard spectrum BDT result at a zenith angle chosen to be 100° for this example.

spectrum. Figure 4.9 shows the result of this procedure for a zenith angle in the horizontal region. It is apparent that there is not a very strong dependence on the cut value if one is not cutting too hard; statistical fluctuations are noticeable. From these plots, the optimal value is chosen and fixed.

The second step, only applied in the upgoing regions, blends in events from the BDTs trained with the softer spectrum, while the discovery potential for an E^{-2} spectrum and a cutoff spectrum[12] of $E^{-2.0} \cdot \exp(-E_{\mathrm{MC}}/10\,\mathrm{TeV})$ is constantly monitored. The test declinations were chosen to be located in the center of the zenith region. It is clear that

Zenith region	tested zenith angle	BDT cut value hard spectrum BDT	BDT cut value soft spectrum BDT
Downgoing	24° and 79°	1.175	-
Horizontal upgoing	100°	1.15	1.275
Vertical upgoing	159°	1.375	1.31

Table 4.8: BDT cut values obtained in the optimization.

the addition of low energy events can slightly degrade the discovery potential for harder spectra. Therefore, a fair compromise has to be found to archive a flexible final sample sensitive to a wide range of spectra. This compromise also involves the requested smooth transition between the zenith regions. The final cut values are shown in Table 4.8.

[12]Different soft spectra and cut off spectra have been tried and resulted in a similar performance [141].

4.3.6 Adjustment of the downgoing event rate

While the two upgoing regions easily connect without additional treatment because of their similarity, the transition at the horizon needs additional adjustment. On the one hand, the downgoing event rate after the BDT cuts is still up to a factor of 5 higher compared to the horizontal upgoing zone. On the other hand background rejection of downgoing particles is mainly based on rejecting lower energies which also leads to sharp transitions in the energy distributions at the horizon.

These issues are solved by a combination of two approaches. First, the horizontal upgoing BDT cut values are gradually raised in the overlap region between a zenith angle from 90° to 85° to smoothly fade out the selection characteristics of the upgoing zone (values in Table B.5 of the appendix). Events cut by the harder values in the overlap can still enter the sample by passing the downgoing BDT. Second, the optimized cut value of the downgoing BDT is adjusted in very fine cosine zenith bins where the rate is too high to match the rate at the horizon and to maintain a constant rate per solid angle. Stronger intervention is only necessary between zenith angles from 90° to 60°, while at lower values the original optimized BDT cut value already produces lower event rates than at the horizon. No large changes in the discovery potential are expected, as it depends only weakly on small changes in the BDT cuts. The adjusted downgoing BDT cut values can be found in Table B.4 of the appendix.

Figure 4.10 shows the final IC79b zenith distribution compared to the original IC79 sample after adjusting cuts. No discontinuities are observable at the transitions between BDT sections. For upgoing cosine zenith values of -0.6 to -0.1, the new selection cuts about 10 % to 15 % harder which is also reflected in the signal efficiency plot in Figure 4.11. Around the horizon a slightly higher event rate is obtained.

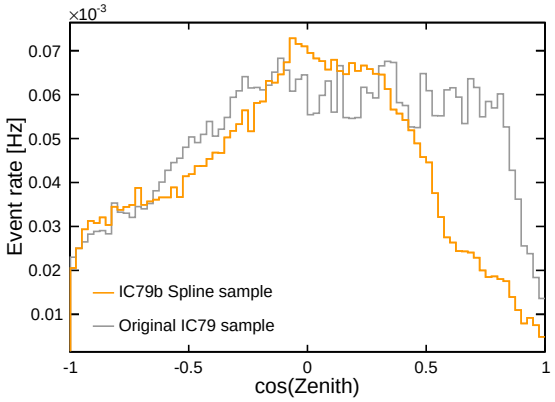


Figure 4.10: Zenith distribution of the final IC79b neutrino sample (orange) and the original IC79 sample (gray). Note that the events for cos(zenith)>0 are dominated by atmospheric muon background.

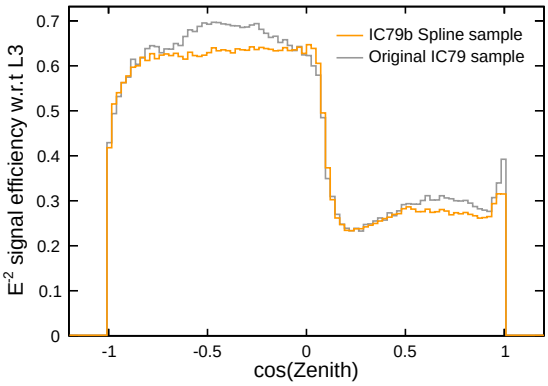


Figure 4.11: E^{-2} signal efficiency with respect to the common starting point at L3 of the original IC79 sample (gray) and the SplineMPE based IC79b sample (orange).

However, the most apparent change is the strong additional background rejection of up to a factor of 3 in the downgoing region above cosine zenith values of 0.4, while the signal efficiency only degrades by 10 %. This improvement is, on the one hand, caused by the optimized BDT configuration and possibly by a slightly better discrimination of atmospheric muon bundles by the Spline based variables on the other hand. The small peak in the signal efficiency at cos(zenith) > 0.9 is a result of the IceTop veto extension to lower energies.

From these plots, an improvement in discovery potential caused by the new selection only is just expected in the downgoing hemisphere by the more efficient rejection of atmospheric muons. Considering the fact that the original IC79 sample already reaches an atmospheric neutrino purity of nearly 100 % in the upgoing region, this is the expected result. However, the better pointing of the SplineMPE reconstruction will improve the sensitivity in the whole sky as shown in Section 4.4.

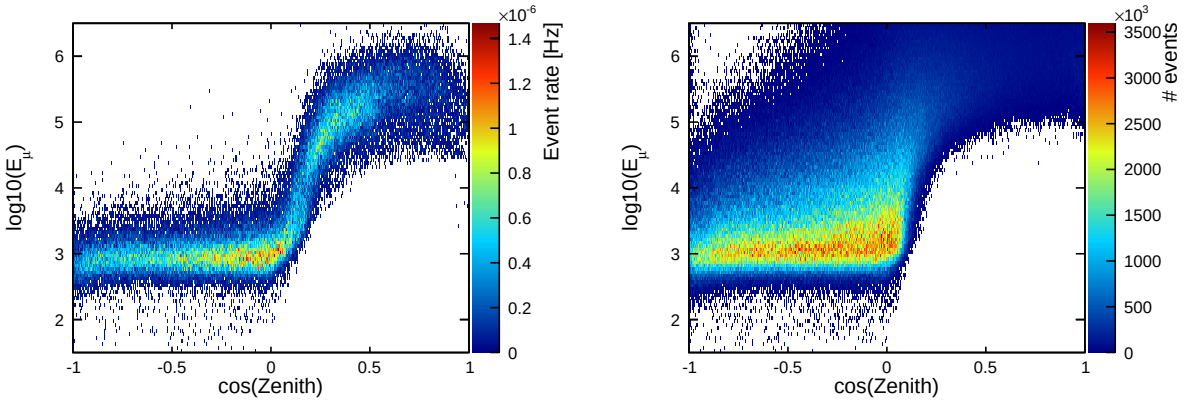


Figure 4.12: Energy-zenith angle distribution of the final IC79b sample for data (left) and a E^{-2} neutrino signal (right). The energy estimator is MuEX in the upgoing region and MuE in the downgoing hemisphere with a linear crossover between a zenith angle of 75° to 90°.

Figure 4.12 presents the final result confirming the seamless transition between the three BDT regions in the energy-zenith plane. Additionally, it is apparent that one has to cut hard on the estimated energy (≈ 20 TeV) to obtain the optimal sensitivity in the southern hemisphere.

Additionally, sanity checks have been performed comparing simulation and data of the IC79b selection. The comparison is shown in Figure B.6 of the appendix and reveals a similar agreement as in the original IC79 sample.

4.4 Final sample properties and summary

In this section I analyze the properties of the new IC79b sample such as angular resolution and discovery potential. Moreover, the optimal configuration of energy and angular accu-

racy estimators for the use in a maximum likelihood ratio analysis, described in Section 5.2, is chosen.

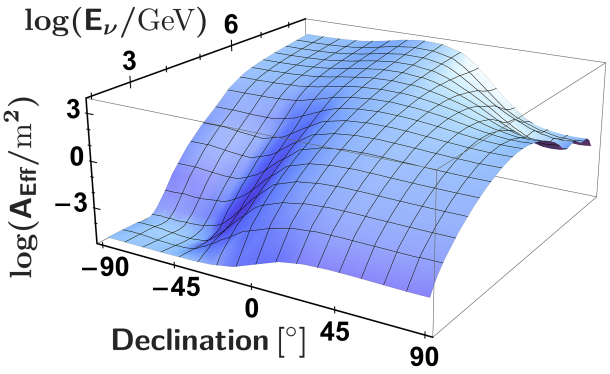


Figure 4.13: IC79b final sample effective area depending on the neutrino energy E_ν and the declination.

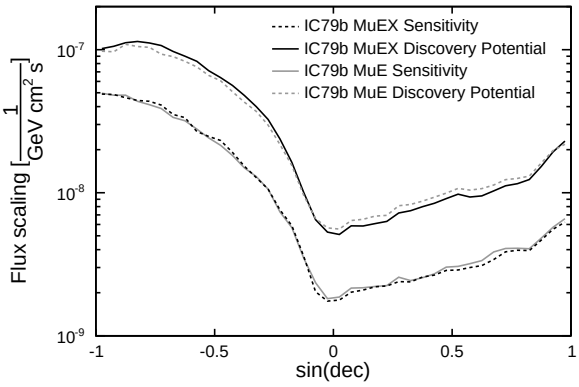


Figure 4.14: Sensitivity and discovery potential given as a scaling factor for an E^{-2} spectrum, comparing MuEX in black and MuE in gray.

The IC79b dataset, defined in the previous sections, consists of 93 842 events – 48 904 from the northern hemisphere and 44 938 from the southern hemisphere. A convenient way to illustrate the response of such a sample to a neutrino flux is given by the effective area A_{Eff}. It connects a neutrino flux with the corresponding expected number of detected neutrinos and is given in Equation 5.7. Figure 4.13 shows A_{Eff} for the IC79b sample. It is easy to see that IceCube is more sensitive in the northern hemisphere at sub PeV energies. Above that energy, neutrinos are absorbed in the earth and the sensitivity is better in the southern hemisphere.

4.4.1 Energy estimator choice

For the point source analyses in IceCube, an energy estimator is used to distinguish soft background and hard signal spectra. It has been shown in Section 4.1.1.1 that MuE assigns lower energies to atmospheric muons than the more advanced energy estimators TE and MuEX. This behavior leads to a better discovery potential in the downgoing region as shown in Figure 4.14 by further separating this background from an astrophysical flux. However, in the upgoing hemisphere this effect is useless due to the lack of the atmospheric muon background. Thus, the energy estimator with the best resolution provides the best results in this case, which is MuEX (see Figure 4.1).

As a result, the sample is configured with MuEX and MuE in the downgoing, respectively upgoing hemisphere. To preserve a seamless transition, a linear crossover is applied in the zenith range from 75° to 90°. Figure 4.12 already shows this optimal energy estimator configuration.

4.4.2 Paraboloid pull correction

An ideal angular accuracy estimator for a point source analysis should be distributed around the true deviation from the Monte Carlo track. This can be inspected by a value called pull, comparing the true error $\angle_{\text{(reco - MC)}}$ and the estimated error by the Paraboloid fit σ_{PB}:

$$\text{pull} = \frac{\angle_{\text{(reco - MC)}}}{\sigma_{\text{PB}}} \,. \tag{4.20}$$

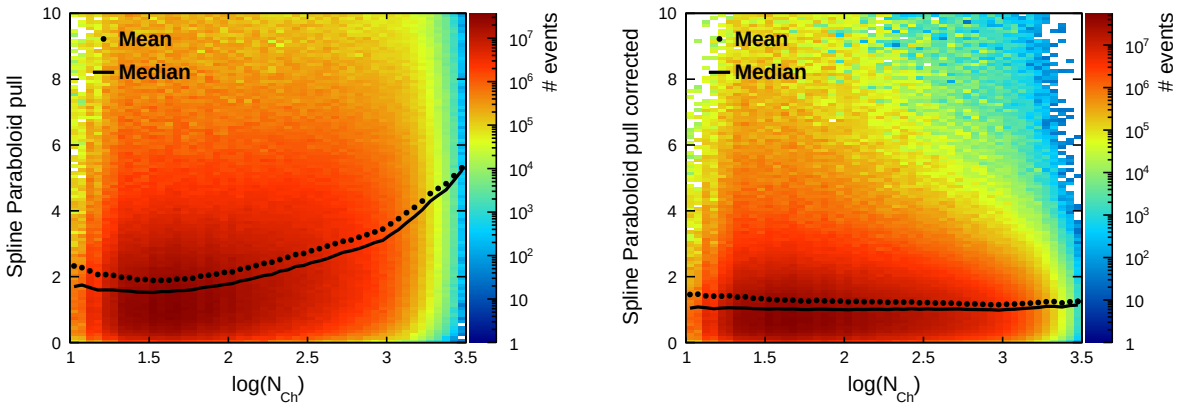


Figure 4.15: Left: Uncorrected SplineMPE Paraboloid pull depending on the number of hit DOMs. Right: Corrected SplineMPE Paraboloid pull based on the adjusted Paraboloid error estimation depending on the number of hit DOMs. The mean (median) per N_{Ch}-bin is shown as dots (line).

The left plot in Figure 4.15 shows the mean and the median of the Paraboloid pull drifting away from the ideal value of 1, indicating an overestimation of the resolution at higher energies. This characteristic is fixed by correcting σ_{PB} with a polynomial depending on N_{Ch}:

$$\begin{aligned}\sigma_{\text{(PB corrected)}} = \sigma_{PB} \cdot (&0.666 + 6.119 \cdot \log(N_{\text{Ch}}) - 9.883 \cdot \log(N_{\text{Ch}})^2 \\ &+ 6.393 \cdot \log(N_{\text{Ch}})^3 - 1.806 \cdot \log(N_{\text{Ch}})^4 + 0.193 \cdot \log(N_{\text{Ch}})^5) \,.\end{aligned} \tag{4.21}$$

The correction is applied to the median to reduce the sensitivity on outliers. The right plot in Figure 4.15 illustrates the pull for $\sigma_{\text{(PB corrected)}}$ and confirms the median to be close to 1. In likelihood point source analyses, this adjusted value is used in the signal PDF.

4.4.3 Angular resolution and sensitivity

The main purpose of the process described in this chapter is the inclusion of the SplineMPE reconstruction into the IC79 data and its use in the final neutrino sample. Figure 4.16

shows the median angular resolution of the new SplineMPE fit and the older PandelMPE fit in the final IC79b sample depending on the neutrino energy. The improvement in resolution still persists at all energies on final level after applying straight cuts and the BDT classification.

The better pointing and selection lead to a better sensitivity throughout the whole sky. Figure 4.17 gives the discovery potential and sensitivity for the original IC79 sample in gray (PandelMPE, MuE) and the new IC79b sample in orange (SplineMPE, MuE and MuEX). The value given is the flux scaling for an E^{-2} spectrum. The improvement depends on the declination and ranges between 20 % and 30 %. A significant improvement could also be achieved for E^{-1} spectra and E^{-3} spectra shown in Figures B.3 and B.4 of the appendix.

The better sensitivity of the IC79b sample, produced in the course of this work, is used in Chapter 5 to analyze the neutrino flux from blazars.

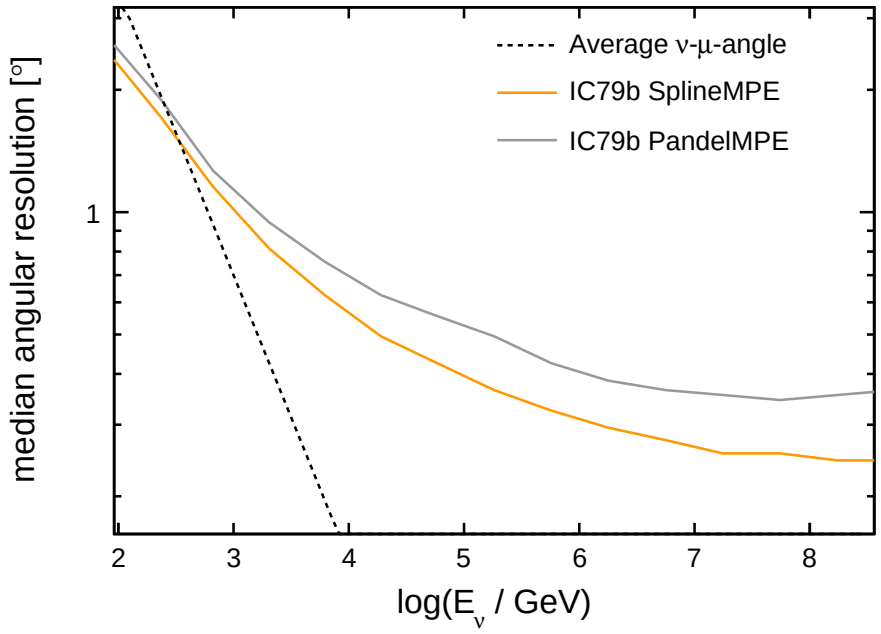


Figure 4.16: SplineMPE angular resolution in the IC79b sample at the final level compared to PandelMPE depending on the neutrino energy.

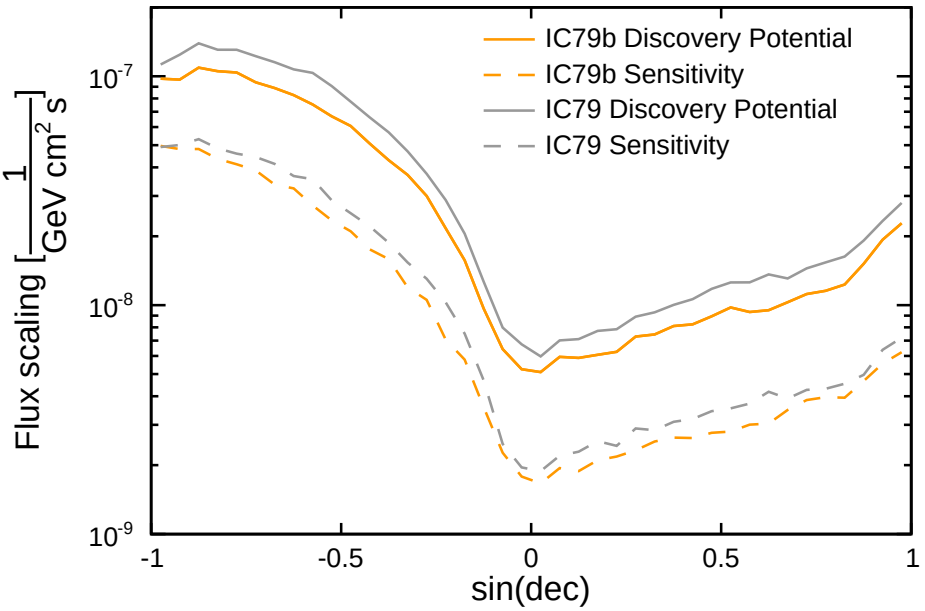


Figure 4.17: E^{-2} sensitivity and discovery potential of the IC79b sample compared to the old IC79 sample depending on the declination.

5

Analysis: Neutrinos from blazars

This chapter describes an unbinned likelihood ratio stacking analysis searching for neutrino emissions from blazars incorporating four years of IceCube data. In a first step, three promising blazar types are identified forming the source catalogs. After the introduction of the analysis method, the selection and the relative weighting of the blazars are specified. The chapter then focuses on the choice of a spectrum cutoff, provides details on the involved detector data and examines the expected sensitivity. After discussing the systematic uncertainties, the analysis results and generic flux limits are presented. Finally, specific flux predictions from three theory papers are investigated and limits for these fluxes are evaluated based on the stacking results.

5.1 Which blazars are neutrino loud?

In hadronic blazar models, pions are created by p-γ (through Δ^+) or p-p interactions resulting in a neutrino flux from decaying charged pions (see Section 2.3.1.1). The Δ^+-resonance or p-p interactions convert a similar amount of the initial proton energy into neutral and charged pions (branching ratio 2:1). IceCube is most sensitive to sources with a high flux in the TeV-ν-range. Thus, the intrinsic γ-luminosity in that range would be a decent indicator for a promising source under the assumption of hadronic acceleration and a significant contribution from π^0-decays to the overall γ-emission. However, the apparent γ-flux is strongly attenuated above 0.1 TeV due to pair production on soft photons from internal sources like the broad line emission region, jet and accretion disk or the extragalactic background light (EBL, Figure 5.1) [29, 75, 142, 143]. As a consequence, TeV-sources can only be detected at low redshifts and the intrinsic TeV-spectrum is unknown.

Therefore, the closest connection to neutrino and cosmic ray production is given by GeV-observations which directly point towards the violent proton acceleration also required for neutrino production at higher energies. Given the high magnetic fields necessary for

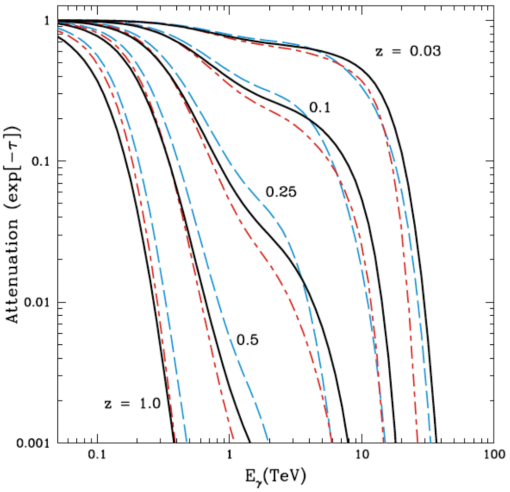


Figure 5.1: γ-attenuation for different models of the extragalactic background light (EBL) as a function of the observed photon energy E_γ for sources at redshifts of 0.03, 0.1, 0.25, 0.5, and 1.0 derived from the optical depth τ [144].

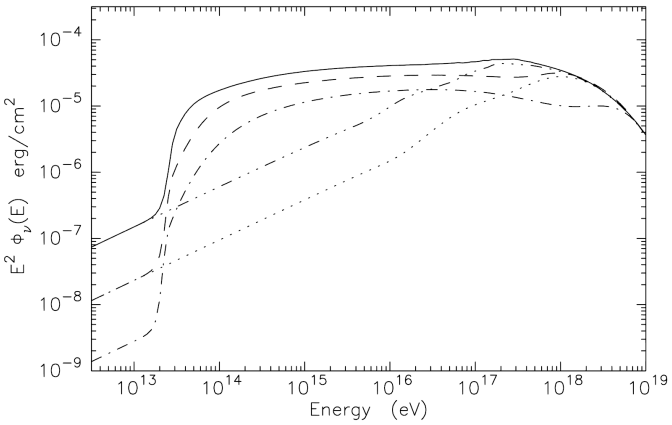


Figure 5.2: The flux Φ_ν, time-integrated over a three week flare of 3C 279, is given for jet Doppler factors of $\delta = 7$ (solid), $\delta = 10$ (dashed) and $\delta = 15$ (dot-dashed). Additionally, the corresponding curves for $\delta = 7$ (triple-dot-dashed) and $\delta = 10$ (dotted) are shown if external radiation fields are neglected [64].

proton acceleration, the high-energy hump in the spectral energy distribution of blazars should be dominated by proton and muon synchrotron radiation [52, 68], complemented by reprocessed TeV-gamma rays or – in case of p-p-models – even dominated by gamma rays from neutral pion decays [75]. Hence, a strong TeV-neutrino emitter is expected to be a powerful GeV-gamma emitter and as a result the neutrino flux is commonly assumed to be proportional or at least correlated with the GeV-emission [145, 146, 147]. Thus, the main selection criterion for all catalogs in this analysis is the photon flux in the Fermi range between 1 GeV to 100 GeV from the second Fermi LAT AGN catalog [55].

The efficiency of the neutrino production from ultra high-energy protons then depends on further characteristics of blazars, like e.g. the magnitude of the target photon field. In the following, three promising blazar classes are identified based on intrinsic properties and specific model predictions. In the end, each group will form its own catalog and each of the three collections will be analyzed separately.

FSRQs: FSRQs are the most luminous blazars and generally found at larger redshifts than BL Lacs. It has been shown that the external radiation field from their broad line emission region could enhance the photopion efficiency by one order of magnitude [64, 143]. As a consequence, these models predict the detection of $\mathcal{O}(1)$ neutrinos per year with IceCube from the brightest sources like e.g. 3C 279. Figure 5.2 shows the predicted ν-flux integrated over the three week flare of 3C 279 in 1996 with and without the external radiation field [64]. Objects with a high intrinsic luminosity ($> 10^{46}\,\mathrm{erg\,s^{-1}}$) are more likely to accelerate protons to ultra high energies [148].

LSP BL Lacs: Without the external photon field from the broad line region the photomeson production in BL Lac objects relies on the synchrotron radiation from the co-accelerated relativistic electrons also forming the lower hump in the spectral energy distribution of blazars. With the generally larger bolometric luminosity of LSPs compared to HSPs, a denser photon field can be concluded for these objects. This translates into a more efficient photopion production and a higher neutrino flux from LSP BL Lacs [68] (Figure 5.3).

Hard γ-spectrum BL Lacs: Neutrinos and photons from photohadronic interactions are produced at energies above $\approx 30\,\mathrm{TeV}$ [64] because of the high pγ interaction threshold of $E_{p\gamma,\mathrm{thr}} = (m_p m_\pi + m_\pi^2)/\epsilon$, where m_p and m_π are the proton and pion mass and ϵ is the target photon energy [75]. Thus, the GeV γ-observation does not allow for a direct estimation of the shape of the TeV neutrino spectrum.

However, p-p interactions are allowed at much lower energies above a threshold of $E_{pp,\mathrm{thr}} = 2m_\pi(1 + m_\pi/m_p) \approx 280\,\mathrm{MeV}$ [75], generating neutrinos already at a fraction of $E_{pp,\mathrm{thr}}$. Under the assumption of a proton power law spectrum and a significant contribution from π_0-decays to the total GeV γ-emission, the TeV ν-flux can be extrapolated. Considering IceCube's energy dependent sensitivity, the blazars with the hardest GeV γ-spectra, which are solely BL Lac objects, could be detectable within p-p models (Figure 5.4) [75].

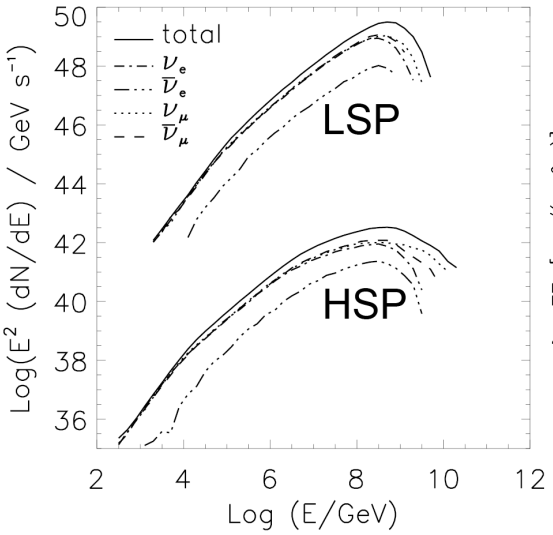


Figure 5.3: Predicted neutrino flux at the source for Mkn 421 (labeled "HSP") and PKS 0716+714 (labeled "LSP") [68].

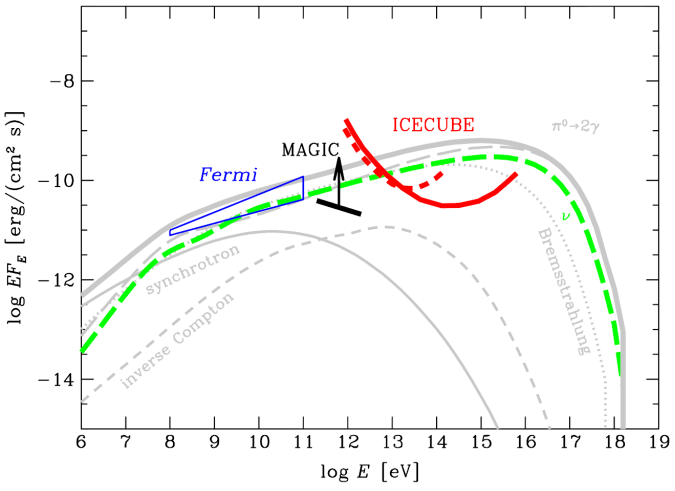


Figure 5.4: Spectra of γ-ray (gray) and neutrino (green) emission of blazar PG 1553 +11 expected in a p-p model in which the GeV emission is dominated by the γ-rays from neutral pion decay. An estimated IceCube sensitivity at two different declinations is shown in red [75].

5.1.1 Model uncertainties

The main problem when searching for neutrinos from blazars is the possibility that totally different gamma emission models could coexist (hybrid approach). Especially the neutrino prediction heavily depends on the relative strength of hadronic and leptonic acceleration. Consequently, it is clear that the connection between gamma ray and neutrino emission fades with rising importance of the inverse Compton scattering. Nevertheless, many pure hadronic blazar models predict fluxes which are just within or even below IceCube's sensitivity. That means that decreasing the hadronic content would shift the emission beyond IceCube's reach. Thus, a discovery oriented analysis within IceCube has to assume a significant hadronic contribution to the gamma flux and a strong connection between gamma and neutrino flux.

Further uncertainties include the shape of the primary proton spectrum, the composition of the radio jet and the intensity and spectral shape of the target photon field. All these values influence the possible resulting neutrino emission. Hence, it is not reasonable to tailor an analysis for one specific model. We therefore decided to follow a general as well as a more specific approach. The analysis will on the one hand assume a very general power law flux and on the other hand, in the more specific approach, an emission following the spectral shape of the gamma emission as motivated by p-p-models.

5.2 Analysis method

The search for point-like neutrino sources is based on a set of muons with direction, energy and time information. An unbinned maximum likelihood ratio test is performed to find a local excess over the atmospheric neutrino flux additionally utilizing energy to distinguish the possibly harder astrophysical spectrum from the softer atmospheric neutrinos (a detailed description can be found in [149] and [109]). Here, the null hypothesis reflects a pure atmospheric data set while the signal hypothesis involves n_s astrophysical source events distributed according to an assumed power law spectrum $E^{-\alpha}$. The test statistic TS is defined as the likelihood ratio of null and signal hypothesis:

$$\mathrm{TS} = -2 \log \left[\frac{\mathcal{L}(n_s = 0)}{\mathcal{L}(\hat{n}_s, \hat{\alpha})} \right] . \tag{5.1}$$

The likelihood of the signal hypothesis $\mathcal{L}$ is optimized with respect to n_s and α[1] delivering TS, $\hat{n}_s$ and $\hat{\alpha}$ as analysis output. Hereby, larger values of TS indicate a lower probability for the null hypothesis. To calculate the significance – or, respectively, the probability that the obtained test statistic value is a result of a pure atmospheric data set (**p-value**) – the overall TS distribution for randomized data has to be known. This can be achieved either by repeated pseudo-runs with events scrambled in azimuth or it can be estimated by a

[1] Here, n_s is restricted to positive values and α may vary between 1 and 4.

χ^2-distribution[2] with two degrees of freedom [150]. Both methods are shown in Figure 5.5 revealing a more conservative (higher) p-value when referring to the χ^2-estimation.

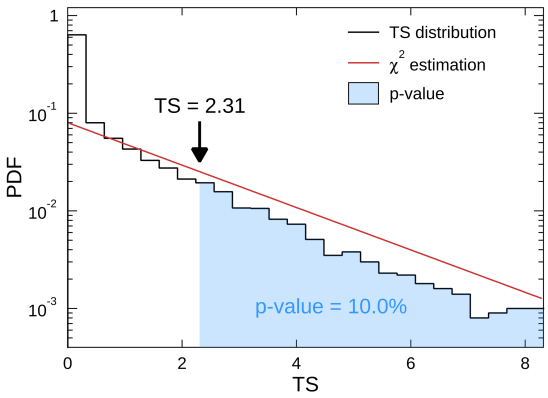


Figure 5.5: P-value evaluation for the LSP BL Lac catalog in the W_{fixed} weighting (see Section 5.3): The TS distribution from scrambled data is shown in black together with the evaluated p-value in blue and the χ^2-approximation in red.

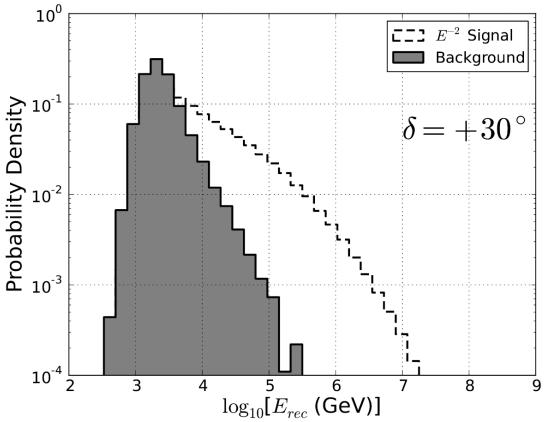


Figure 5.6: Energy PDF vs. reconstructed energy for atmospheric background and an E^{-2}-signal at a declination of 30° [109].

In IceCube, several datasets for different years have to be combined into one global likelihood value $\mathcal{L}$ by multiplying the likelihoods for the particular years j (Equation 5.2). While the power law index is assumed to be identical within all sets ($\alpha_j = \alpha$), the total number of signal events n_s is split. The expected fraction of source events for a given data set n_s^j is derived from simulation and depends on the effective area of each data set. $\mathcal{L}$ can then be written as

$$\mathcal{L}(n_s, \alpha) = \prod_j \mathcal{L}^j(n_s^j, \alpha) = \prod_j \prod_{i \in j} \left[\frac{n_s^j}{n_{\text{tot}}^j} \cdot \mathcal{S}_i^j + \left(1 - \frac{n_s^j}{n_{\text{tot}}^j}\right) \cdot \mathcal{B}_i^j \right] \tag{5.2}$$

where $\mathcal{S}_i^j$ and $\mathcal{B}_i^j$ are the signal and background PDFs for event i in the data set j and n_{tot}^j is the total number of events in the data set j. $\mathcal{B}_i^j$ consists of the event density per unit solid angle $B_i^j(\delta_i)$ which is calculated per zenith band as well as the energy PDF $\mathcal{E}_i^j(E_i, \delta_j)$ giving the probability to observe an atmospheric particle with energy E_i at the declination δ_i:

$$\mathcal{B}_i^j = B_i^j(\delta_i) \cdot \mathcal{E}_i^j(E_i, \delta_i)\,. \tag{5.3}$$

[2]The χ^2-estimation relies on the validity of Wilks's theorem which holds if the null hypothesis is a special case of the signal hypothesis. This is the case here for $n_s = 0$.

The signal PDF $\mathcal{S}_i^j$ can be written as

$$\mathcal{S}_i^j = S_i^j(|\vec{x}_i - \vec{x}_s|, \sigma_i) \cdot \mathcal{E}_i^j(E_i, \delta_i, \alpha)\,, \tag{5.4}$$

where again $\mathcal{E}_i^j(E_i, \delta_j, \alpha)$ is the energy PDF expressing the probability to observe a muon with energy E_i at the declination δ_i for a power law spectrum with index α. An example for the energy PDFs of background and signal is shown in Figure 5.6. The spatial term $S_i^j(|\vec{x}_i - \vec{x}_s|, \sigma_i)$ reflects the probability that a particle with direction $\vec{x}_i$ belongs to a source at $\vec{x}_s$ and is modeled by a Gauss function:

$$S_i^j = \frac{1}{2\pi\sigma_i^2} \cdot e^{-\frac{|\vec{x}_i - \vec{x}_s|^2}{2\sigma_i^2}}\,. \tag{5.5}$$

An angular resolution estimator, like the Paraboloid fit introduced in Section 4.1.2, delivers the standard deviation σ_i. The sensitivity on a uniform group of astrophysical objects can be enhanced by performing a stacking analysis testing the summarized flux from a catalog of potential sources [122]. In case of a stacking analysis, the signal PDF is modified towards a weighted sum over M sources:

$$\mathcal{S}_i \rightarrow \mathcal{S}_i^{\text{Stack}} = \frac{\sum_{k=1}^{M} W^k R^k(\alpha, \delta_k) \cdot S_i^k(|\vec{x}_i - \vec{x}_k|, \sigma_i) \cdot \mathcal{E}_i(E_i, \delta_i, \alpha)}{\sum_{k=1}^{M} W^k R^k(\alpha, \delta_k)}\,. \tag{5.6}$$

Every source is weighted by a detector term $R^k(\alpha, \delta_k)$, representing the relative detection efficiency at the source's declination δ_k for a given power law spectrum with index α. Additionally, a theoretical weight W^k can be used to scale a source in the likelihood according to the expected neutrino output.

5.2.1 Discovery potential, sensitivity and limits

For the calculation of discovery potential, sensitivity and flux limits the detector data is scrambled in azimuth angle and used as background estimate. Simulated signal events are then injected following a predefined flux function and distributed among the different years following the method used in the likelihood description. In the stacking case, every source may inject a different flux. The analysis is then performed multiple times with different numbers of source events and the test statistic and the p-value is calculated. The result of the evaluation is the normalization constant for the total injected flux complying with the following definitions:

- **Discovery potential:** Flux which yields a p-value smaller than 5σ (2.87×10^{-7}) in 50% of the runs.
- **Sensitivity:** Flux which creates a p-value smaller than 0.5 in 90% of the runs.
- **90% confidence level (CL) limit:** Flux which produces a p-value smaller than the p-value from the analysis with real unscrambled data in 90% of the runs.

For computational reasons, the p-value is estimated on the basis of the χ^2-function in the intermediate calculations while the final p-value of the analysis is evaluated accurately from the TS distribution obtained from scrambled data. The discovery potential, sensitivity and the limits are always given as a scaling factor for the injected test flux function.

5.3 Source weighting

As stated in Section 5.1, a close connection between γ-flux and ν-flux is generally assumed. Therefore the theoretical source weight W^k from Equation 5.6 is chosen to be the integrated γ-flux $F_\gamma = \int_{1\,\mathrm{GeV}}^{100\,\mathrm{GeV}} \Phi_\gamma dE$ measured by Fermi LAT [55] and defined as photons per centimeter and per second. For the detector weight $R^k(\alpha, \delta_k)$ two different weighting models are adopted:

- **variable detector weight ($W_{\mathbf{variable}}$):** The spectral index α in the detector weight is identical to the global index fitted in the likelihood optimization. As the detector weight $R^k(\alpha, \delta_k)$ strongly depends on this spectral index, the source weight is variable during the likelihood optimization. Aside from the normalization W^k, this approach assumes the same power law flux from all sources and ignores the shape of the gamma spectrum.

- **fixed detector weight ($W_{\mathbf{fixed}}$):** Motivated by the possible extrapolation of the TeV neutrino flux from the GeV γ-emission in p-p models (Section 5.1), the spectral index α in the detector weight is defined as the γ-ray power law index evaluated from the Fermi LAT data. As a result, the source weight is fixed during the likelihood optimization and hard γ-emitters enter with a higher weight[3].

Due to the large uncertainties and the variety of blazar models each of the three blazar groups discussed in Section 5.1 is analyzed with both weighting schemes resulting in a total of 6 analyses.

5.4 Blazar selection

The stacking method with weighted sources used in this work could in principle be performed on all Fermi blazars within one of the groups defined in Section 5.1 without a penalty on sensitivity. Nevertheless, it is convenient to cut the catalog to a small and clear selection by omitting sources with negligible weight.

Regarding the variable detector weight approach W_{variable}, the first selection criteria is the integrated γ-flux F_γ also used as theoretical source weight. To reflect the fixed detector weight approach W_{fixed}, the expected detected number of neutrinos N_ν has to be calculated as second selection criteria for every source under the assumption that the ν- and γ-flux show the same spectral shape:

[3] This is a direct result of IceCube's better sensitivity at energies above 1 TeV, as shown in Figure 4.13

$$dN_\nu = c_1 \cdot \Phi_\gamma \cdot A_{\mathrm{Eff}}(\delta, E) \cdot T \cdot dE\,. \tag{5.7}$$

Here, c_1 is a proportionality constant, Φ_γ the differential gamma flux in units of $\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, A_{Eff} the effective area, δ the declination, E the energy and T the measurement duration. As N_ν is only used as a relative number for source selection the constants can be chosen arbitrarily to be $c_1 = 1$ and $T = 1\,\mathrm{yr}$. The differential gamma flux can be written as

$$\Phi_\gamma = F_\gamma \cdot \frac{(E/\mathrm{GeV})^{-\alpha_\gamma}}{\int_{1\ \mathrm{GeV}}^{100\ \mathrm{GeV}} (E/\mathrm{GeV})^{-\alpha_\gamma}\, dE}\,, \tag{5.8}$$

with the integrated Fermi flux F_γ and the Fermi power law index α_γ. A source enters the catalog by meeting the γ-flux or the N_ν requirements listed in Table 5.1. Additional restrictions apply for hard γ-spectrum BL Lacs ($\alpha_\gamma < 2.3$) and for FSRQs (Luminosity(0.1 GeV to 100 GeV)$> 10^{46}\,\mathrm{erg\,s}^{-1}$). The luminosity is defined as [151]

$$L_\gamma = 4\pi d_L^2 \frac{S(0.1\,\mathrm{GeV}\ to\ 100\,\mathrm{GeV})}{(1+z)^{2-\alpha_\gamma}}\,, \tag{5.9}$$

where $S(0.1\,\mathrm{GeV}\ to\ 100\,\mathrm{GeV})$ is the energy flux taken from Fermi data between the given energies [55], z is the redshift, α_γ is the photon spectral index and d_L is the luminosity distance[4].

Type	$\boldsymbol{N_\nu}$	$\boldsymbol{F_\gamma}$	$\boldsymbol{L_\gamma}$	$\boldsymbol{\alpha_\gamma}$
FSRQs	> 1	> 8	$>10^{46}$	-
LSP BL Lacs	> 1	> 3	-	-
hard γ-spectrum BL Lacs	> 50	> 5	-	< 2.3

Table 5.1: Blazar catalog selection criteria. F_γ is given in units of $10^{-9}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and L_γ in units of $\mathrm{erg\,s}^{-1}$.

The cut values in Table 5.1, chosen to discard sources with negligible weights in both weighting models, result in 33 FSRQs, 27 LSP BL Lacs and 37 hard γ-spectrum BL Lacs. The selected sources and their properties are listed in Tables B.6, B.7 and B.8 of the appendix.

5.5 Dominant sources and cutoff study

A likelihood stacking is characterized by the particular source weights, defining their contribution to the analysis. This section discusses the dominant sources within the likelihood

[4]For the calculation of the luminosity distance d_L, $H_0 = 0.673\,\mathrm{km\,s}^{-1}\,\mathrm{Mpc}^{-1}$, $\Omega_\Lambda = 0.685$ and $\Omega_M = 0.315$ [16] are used for the cosmological parameters.

optimization, which are listed in Table 5.2. The objects contributing the most span from weights of 8% to 19% in W_{variable} at an extreme spectral index $\alpha = 1.5$ and vary around 30% for W_{fixed}. As a first conclusion it can be stated, that no source weight diverges at extremely hard indexes within W_{variable}. Nevertheless, if the highest source weights are a direct consequence of the assumed spectral shape in the detector weights $R^k(\alpha, \delta_k)$ (Equation 5.6), a theoretically motivated spectrum cutoff might attenuate these sources, and further equalize the catalogs.

Catalog	$\boldsymbol{W}_{\text{variable}}$ ($\alpha = 1.5$)	$\boldsymbol{W}_{\text{fixed}}$
FSRQ	3C 454.3: 19 %	4C +55.17: 30 %
LSP BL Lacs	AO 0235+164: 15 %	PKS 1717+177: 31 %
Hard γ-spectrum BL Lacs	AO 0235+164: 8 %	PG 1553+113: 32 %

Table 5.2: The dominant sources in the likelihood stacking for W_{variable} at a spectral index $\alpha = 1.5$ and W_{fixed} are shown with their relative, normalized weight.

A neutrino cutoff can be derived from the primary proton cutoff. It is estimated to be around $\approx 10^9$ GeV from reasonable acceleration scenarios. As a result, the exponential neutrino cutoff can be assumed to be about a factor of 10 below [75]. Within IceCube, the effect of such a high neutrino cutoff at 10^8 GeV is expected to be negligible, hence additionally a lower boundary at 10^6 GeV was tested. Both cutoff variants were applied to the injected spectrum and to the detector weights, and sensitivity and discovery potential were evaluated.

The likelihood source weights were monitored and revealed only negligible changes within a few percent due to the cutoffs. Thus, one has to conclude that a cutoff cannot attenuate the highest weights. The effect on the sensitivities was also negligible for a cutoff at 10^8 GeV. When using 10^6 GeV, the sensitivity degraded by a factor of 1.5 in all cases except for the hard γ-spectrum BL Lacs in the W_{fixed} weighting. Because of the hard assumed spectra, this scenario is affected the most and the sensitivity degrades by a factor of 2.

As a result of the observed small impact, the exponential cutoff at 10^8 GeV, as motivated theoretically, is used within this analysis for all injected fluxes and the detector weight.

5.6 Analysis data

The analysis covers four years of IceCube data including three years from the partially completed detector with 40, 59 and 79 strings (IC40, IC59, IC79b) and one data taking period with the finalized experiment (IC86). The data sets from IC40 and IC59 are described in detail in [109, 122] while the IC86 sample is specified in [120]. The reprocessed IC79b sample discussed in Chapter 4 is used for the 79 string configuration.

In IC40, the event selection is based on a set of manually optimized cuts on a collection of heavily studied quality variables complemented by a declination dependent energy cut

Configuration	Livetime [days]	Total events	upgoing	downgoing
IC40	376	36 900	14 121	22 779
IC59	348	107 569	43 339	64 230
IC79b	316	93 842	48 904	44 938
IC86	333	138 322	69 227	69 095
Total	1373	376 633	175 591	201 042

Table 5.3: Analysis data statistics

for the southern sky. In IC59 multivariate signal classification with boosted decision trees (Section 4.3) has been introduced for the first time to improve the selection efficiency and sensitivity. Additionally, also the IceTop veto pictured in Section 4.1.3 has been utilized for atmospheric muon suppression for the first time. The main upgrade for the IC79b and IC86 data compared to previous years is the improved angular muon reconstruction developed in Chapter 3. The number of events in a data set depends on the detector configuration and on the background rejection cuts.

The resulting data collection, summarized in Table 5.3, comprises a livetime of 1373 days, a total of 376 633 events with 175 591 (201 042) originating from the northern (southern) sky. The point source sensitivity reaches $0.5 \times 10^{-9}\,\mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at the horizon for an E^{-2} neutrino spectrum.

5.7 Expected performance

The sensitivity and the discovery potential are calculated according to the method specified in Section 5.2.1. For the variable source weighting W_{variable}, every source i injects an E^{-2}-flux weighted by the source's integrated gamma-flux F_γ^i with an exponential cutoff at $10^8\,\mathrm{GeV}$

$$\Phi_{\mathrm{E2}}^i = \frac{F_\gamma^i}{\sum_j^{n_{\text{sources}}} F_\gamma^j} \cdot (E_\nu/\mathrm{GeV})^{-2} \cdot e^{-\frac{E_\nu}{10^8\,\mathrm{GeV}}} \cdot \mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,. \tag{5.10}$$

Due to the weight normalization, the experimentally determined sensitivities refer to the total E^{-2} flux from the whole catalog given by

$$\Phi_{\mathrm{E2}} = \sum_i^{n_{\text{sources}}} \Phi_{\mathrm{E2}}^i = (E_\nu/\mathrm{GeV})^{-2} \cdot e^{-\frac{E_\nu}{10^8\,\mathrm{GeV}}} \cdot \mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,. \tag{5.11}$$

The corresponding injection for the fixed source weighting W_{fixed} additionally includes a source's gamma spectral index α_γ

$$\Phi^i_{\mathrm{MF}} = \frac{F^i_\gamma}{\sum_j^{n_{\mathrm{sources}}} F^j_\gamma} \cdot (E_\nu/\mathrm{GeV})^{-\alpha_{\gamma,i}} \cdot e^{-\frac{E_\nu}{10^8\,\mathrm{GeV}}} \cdot \mathrm{GeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,. \tag{5.12}$$

In this case, the resulting total flux is a weighted, normalized sum of differently shaped spectra indicated by the abbreviation Φ_{MF} (model flux):

$$\Phi_{\mathrm{MF}} = \sum_i^{n_{\mathrm{sources}}} \Phi^i_{\mathrm{MF}}\,. \tag{5.13}$$

Table 5.4 summarizes the evaluated expected sensitivities. It is obvious that the E^{-2} flux normalizations for the W_{variable} weighting are arranged within a narrow range while the values for the model function Φ_{MF} for the W_{fixed} weighting heavily depend on the average spectral index of the catalog. As a result, the FSRQs and LSP BL Lacs with their spectra being softer than E^{-2} require a higher flux normalization than the hard γ-spectrum BL Lacs for the Φ_{MF} injection.

Weighting	**Sensitivity**	**Discovery potential**
	FSRQs	
W_{variable}	$2.8 \cdot \Phi_{\mathrm{E2}}$	$9.7 \cdot \Phi_{\mathrm{E2}}$
W_{fixed}	$22 \cdot \Phi_{\mathrm{MF}}$	$83 \cdot \Phi_{\mathrm{MF}}$
	LSP BL Lacs	
W_{variable}	$4 \cdot \Phi_{\mathrm{E2}}$	$16 \cdot \Phi_{\mathrm{E2}}$
W_{fixed}	$6.7 \cdot \Phi_{\mathrm{MF}}$	$27 \cdot \Phi_{\mathrm{MF}}$
	hard γ-spectrum BL Lacs	
W_{variable}	$3.7 \cdot \Phi_{\mathrm{E2}}$	$15 \cdot \Phi_{\mathrm{E2}}$
W_{fixed}	$0.19 \cdot \Phi_{\mathrm{MF}}$	$0.65 \cdot \Phi_{\mathrm{MF}}$

Table 5.4: Blazar stacking sensitivity and discovery potential. The flux normalizations are given in the order of 10^{-9} relative to the total injected catalog flux. Φ_{E2} refers to the E^{-2} flux function from Equation 5.11 and Φ_{MF} is the total injected model flux from Equation 5.13.

5.8 Systematic uncertainties

The background estimation for the analyses presented in this work is achieved by scrambling the final data samples in azimuth angle. As a result, the evaluated p-values are neither affected by possible systematic uncertainties from the atmospheric muon and neutrino simulation nor from the modeled detector response. However, the limit and sensitivity calculations rely on the injection of simulated signal events which comprise uncertainties from

different sources. Additionally, the energy PDF in the likelihood analysis method is derived from simulated data. In [109], detailed studies were conducted to measure the total effect on IceCube limits by performing the whole simulation, selection and analysis chain with data sets based on deviating fundamental simulation values.

The main contributions to the total error originate from the absolute DOM quantum efficiency and the underlying ice model. Using data sets with the DOM efficiency modified by $\pm 10\%$ caused limit and sensitivity deviations of $+6\%$ and -7%, respectively. The adjustment of absorption and scattering values by $\pm 10\%$ revealed an ice model uncertainty of $+5\%$ and -8%, respectively. Other contributions, like random coincidences in the IceTop veto and the relative DOM efficiency, showed a negligible impact on the sensitivity calculation. The combined maximal uncertainty from muon energy losses, neutrino cross sections and the rock density is stated as $\pm 4\%$ in [122].

Source	**Effect on the sensitivity**
Absolute DOM efficiency	$+6\%/-7\%$
Ice model (Pandel)	$+5\%/-8\%$
Ice model (Splines)	$+16\%/-8\%$
IceTop veto random coincidences	$< 1\%$
Relative DOM efficiency	$< 1\%$
Muon energy losses, neutrino-nucleon cross-section and rock density	$+4\%/-4\%$

Table 5.5: Systematic uncertainties of IceCube point source analyses.

These numbers are valid for the analysis presented here, with one exception: in [109], the new SplineMPE reconstruction introduced in Section 3 was not yet developed. It is shown in [120] that this new method is more sensitive to ice model variations and the effect on the limits and sensitivity increases to $+16\%$ and -8%, respectively. As SplineMPE is used in two of the four data taking periods, the impact is averaged. Hence, the quadratic sum of the single uncertainties, listed in Table 5.5, results in a total uncertainty of 24% $(+13\%, -11\%)$ for the limits and sensitivities presented in this work.

5.9 Results of the blazar stacking

The 6 stacking analyses performed in the course of this work are presented in Table 5.6. Five over-fluctuations from the atmospheric background have been observed, all of them compatible with the null hypothesis. Here, all p-values are calculated from the TS distribution as displayed in Figure 5.5. The lowest p-value of 10% originates from the LSP BL Lac catalog with the W_{fixed} source weighting where the rather soft spectral index of 3.6 indicates an excess caused by low energy events. The hardest power law index, $\hat{\alpha} = 2.45$, shows up in the FSRQ catalog searches.

Weighting	TS	$\hat{n}_s$	$\hat{\alpha}$	p-value	$\Phi^{90\%}_{\nu_\mu+\bar{\nu}_\mu}$
	FSRQs				
W_{variable}	0.51	9.83	2.45	31.0 %	$3.46 \cdot \Phi_{\text{E2}}$
W_{fixed}	1.05	15.39	2.75	19.1 %	$34.3 \cdot \Phi_{\text{MF}}$
	LSP BL Lacs				
W_{variable}	0.30	11.89	3.25	38.4 %	$5.24 \cdot \Phi_{\text{E2}}$
W_{fixed}	2.31	21.81	3.59	10.0 %	$13.5 \cdot \Phi_{\text{MF}}$
	hard γ-spectrum BL Lacs				
W_{variable}	0	0	-	-	$3.73 \cdot \Phi_{\text{E2}}$
W_{fixed}	0.77	17.48	3.95	29.0 %	$0.284 \cdot \Phi_{\text{MF}}$

Table 5.6: Blazar stacking results: Shown are the test statistic TS, the number of signal events $\hat{n}_s$ and the power law index $\hat{\alpha}$ giving the maximum likelihood, the p-value derived from the TS distribution, and the 90% confidence level upper limit $\Phi^{90\%}_{\nu_\mu+\bar{\nu}_\mu}$ for the combined $\nu_\mu + \bar{\nu}_\mu$ flux in orders of 10^{-9}.

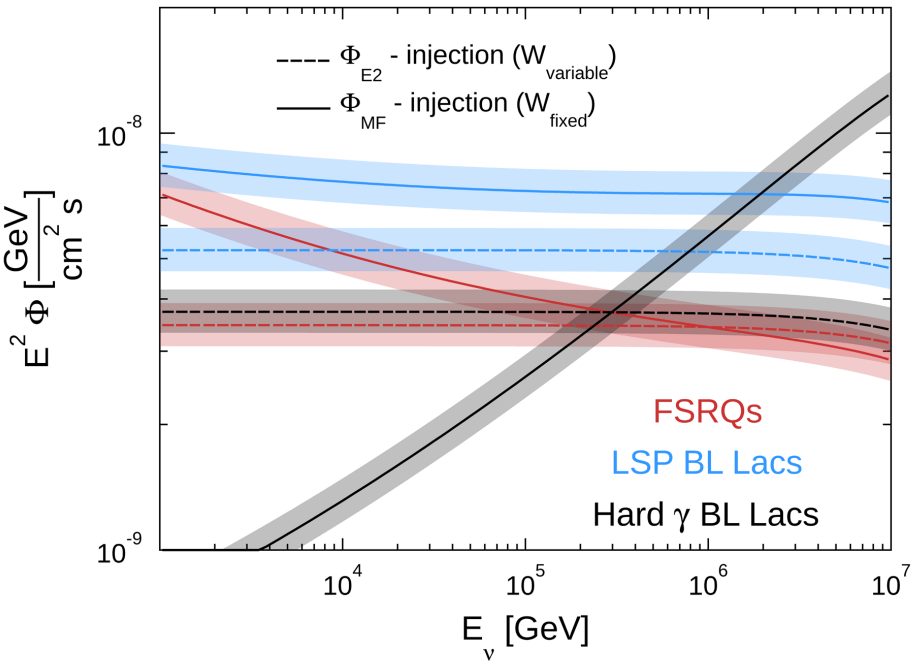


Figure 5.7: 90% CL limits are shown for the injected E^{-2}-flux Φ_{E2} in the W_{variable} weighting (dashed) and the model function flux Φ_{MF} in the W_{fixed} weighting (solid) for the total flux of each catalog. The error bands illustrate IceCube's systematic uncertainty.

All catalogs show higher TS and $\hat{n}_s$ values and a lower p-value for the fixed source weighting W_{fixed} which prefers hard gamma emitters. However, giving the high probability that this is a random coincidence and the compatibility with background, it is not possible to draw further conclusions.

All limits from Table 5.6 are visualized in Figure 5.7, especially showing the shape of the injected model function flux Φ_{MF} for all catalogs. It is evident that – as a consequence of the γ power law index – hard γ-spectrum BL Lac objects inject a considerable harder flux than E^{-2} for the W_{fixed} weighting while the FSRQ and LSP BL Lac fluxes are assumed to be slightly softer.

5.10 Model specific limits

The three different blazar groups analyzed in this work have been motivated by theoretical calculations for specific sources of each object type. The expected neutrino flux from the FSRQ 3C 279 is given in [64] (Figure 5.2), the neutrino emission from the LSP BL Lac PKS 0716[5] is predicted in [68] (Figure 5.3) and the hard γ-spectrum BL Lac PG 1553 is modeled in [75] (Figure 5.4). These predicted muon neutrino fluxes are shown as dashed lines in the top plot in Figure 5.8. One observes that the external radiation field from the broad line region of the FSRQ (red) extends the photo pion production and neutrino emission towards lower energies compared to the LSP BL Lac (blue). The lowest energies are expected from the very optimistic model of the hard γ-spectrum BL Lac in the light of p-p interactions (black). The goal is to evaluate limits for these specific flux predictions.

In a first step, the three modeled objects were analyzed separately in a non-stacking point source search. The obtained p-values, all well compatible with background, can be used to evaluate limits for these flux expectations as defined in Section 5.2.1. The resulting limits are shown as solid bands in the top plot in Figure 5.8, with the systematic errors discussed in Section 5.8. Figure 5.8 demonstrates that only the very optimistic model for PG 1553 can be rejected, while the upper limits for 3C 279 and PKS 0716 are still above the flux prediction. The loose limit for PKS 0716 is a result of its declination at 71° where the predicted high-energy neutrinos are absorbed in the earth.

In the next step, these limit calculations are applied to the stacking results produced in this work. Therefore, the total expected flux from a whole source catalog has to be evaluated from the predictions for 3C 279[6], PKS 0716 and PG 1553, respectively, which are used as a template. This is possible under the assumption that the neutrino flux is proportional to the gamma flux.

These total predicted catalog fluxes are displayed as dashed lines in the bottom plot in Figure 5.8. The corresponding upper limits can be calculated with the stacking results in Table 5.6 and are drawn as solid bands, also indicating the systematic uncertainties. Ad-

[5]PKS 0716 was classified as LSP at the time the paper was published while it is categorized as ISP today.

[6]The flux for 3C 279 in [64] is given for the flaring period in 1996. For this work, the average flux is calculated using a factor 10 lower flux in quiescent state [152], a 10% duty cycle [64] and the prediction for a jet Doppler factor of 7.

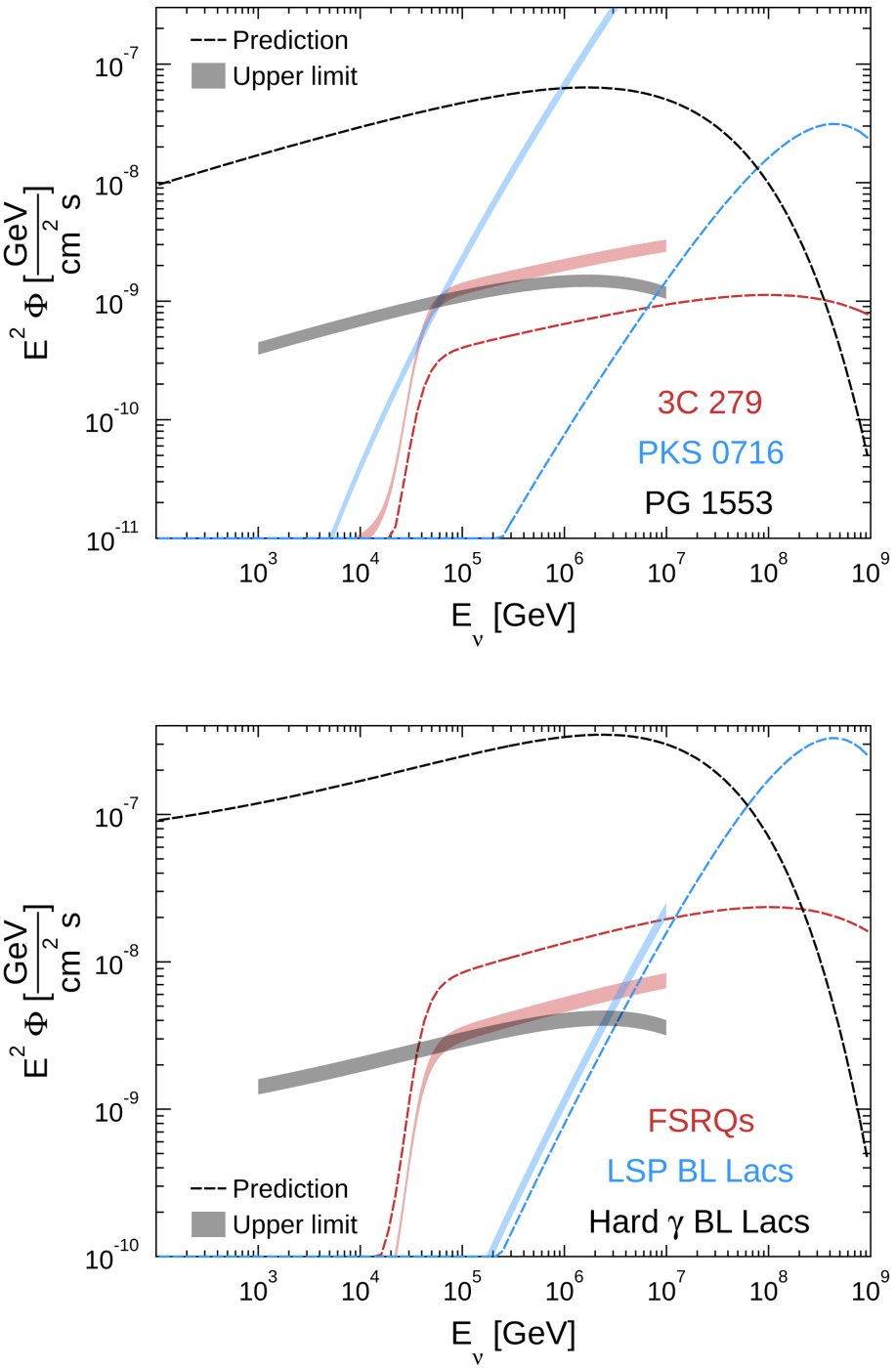


Figure 5.8: In the top plot, muon neutrino flux predictions are shown as dashed lines for 3C 279 [64] (red), PKS 0716 [68] (blue) and PG 1553 [75] (black). The lower plot shows the muon neutrino flux predictions from the FSRQ (red), LSP BL Lac (blue) and hard γ-spectrum BL Lac (black) catalogs evaluated from the single source models at the top. In both cases, the corresponding 90% CL upper limits are drawn as solid bands also indicating IceCube's systematic uncertainty.

ditionally, Figure 5.8 demonstrates how a stacking analysis increases IceCube's sensitivity relative to an expected flux compared to single source analyses.

Of course, the optimistic p-p model upper limit (black), already rejected from the single source limit, is again below the prediction (by a factor of ≈ 100). Contrary to the single source analysis, also the hadronic FSRQ flux calculation is challenged by the stacking limit, while the LSP BL Lac limit is only a factor of ≈ 1.5 above the prediction.

It is obvious that such a total flux estimation as performed here will involve large uncertainties. First, it is not clear how good the proportionality assumption reflects the real ν-γ-connection. Second, even sources within one class can show different emission profiles which interferes with the concept of a scalable template flux.

As a consequence, only the flux prediction for p-p interactions in PG 1553 from [75] can be clearly rejected. The FSRQ model from [64] is only challenged under the assumptions made above and could be saved by a larger jet Doppler factor which would decrease the neutrino output. Finally, the hadronic emission model for LSP BL Lacs [68] is – even for a stacking analysis – still slightly below IceCube's sensitivity but could be reachable within a few years of data taking.

6
Summary

The detection of a diffuse astrophysical neutrino flux by IceCube was a milestone in the era of high-energy neutrino astronomy. Obviously, the next step is the identification of single neutrino sources in the sky. To this day, all point source analyses were compatible with the atmospheric background hypothesis. Hence, besides the construction of new detectors and the accumulation of additional data, the existing analysis methods have to be refined.

The technical part of the thesis addresses this issue by the development of an improved angular reconstruction for muons. To this end, the muon light production is simulated and the emitted photons are propagated on graphical processing units following a precise model of the scattering and absorption in the Antarctic ice. From this simulation, probability density functions for the light distribution in space and time are stored in very finely binned tables which are in turn fit by multidimensional spline surfaces. Using these distributions for the reconstruction of muon tracks improves the angular resolution significantly throughout all energies above IceCube's lower threshold of $\approx 0.1\,\mathrm{TeV}$. Additional modifications, specifically incorporating an accurate noise model, sophisticated photon hit cleaning using a Kolmogorov-Smirnov test and refined treatment of timing uncertainties, further improve the accuracy. At the highest energies around 100 PeV, the median angular resolution on an intermediate data selection level reaches 0.4°, whereas old methods achieved 0.7°. At the final data selection level used in point source analyses, the new method even yields resolutions down to 0.25°, a factor of 4 better than anticipated in the original design of IceCube.

The effect of the new reconstruction on point source analyses is demonstrated by the reprocessing of the IceCube data from 2010, where 79 strings were operating. The more accurate pointing improves the discovery potential for point source searches by 20 % to 30 % depending on the declination. At this time, the method developed in this thesis is the standard muon reconstruction for point source analyses in IceCube.

6. SUMMARY

An important task in detector operations is the monitoring of the data quality. In the course of this thesis, a web-interface was developed for this purpose. A wide range of monitoring quantities are delivered via satellite by the monitoring backbone and processed and stored in a MySQL database. General status information, trigger and filter rates as well as various quantities for all IceCube DOMs can be displayed in a web browser for all data-taking periods. IceCube monitoring shift takers analyze the information and can report potential issues via the interactive interface. This feedback is valuable input for the evaluation of the quality of a data taking period.

The sensitivity of point source analyses in IceCube can be enhanced by analyzing the flux from a uniform group of potential neutrino sources instead of single locations in the sky. A maximum likelihood stacking approach was used in this work to search for neutrino emissions from three blazar groups. These are FSRQs, LSP BL Lacs and hard γ-spectrum BL Lacs. All results are compatible with the background-only hypothesis and flux limits were evaluated. It was shown that BL Lac models predicting a γ-emission dominated by neutral pion decays from proton proton collisions can be rejected at least at a 90% confidence level. Moreover, while LSP BL Lac models are still beyond IceCube's sensitivity, hadronic FSRQ models based on proton gamma interactions are challenged under certain assumptions. A publication incorporating the results of this thesis is under review [120].

We know there are astrophysical neutrinos, but their origin remains an open question – for today.

Appendix A

Data quality monitoring

The monitoring of the data quality and the smooth operation of all detector components is crucial in experiments like IceCube. As usually only events from flawless detector operation periods can be used in analyses, problems have to be identified and solved promptly to maximize the available data. An independent monitoring system collects various quality parameters, status information and data rates and transmits the result to the northern hemisphere via satellite. Part of this thesis was the development of an interactive web-interface, which displays the information and allows interactive feedback.

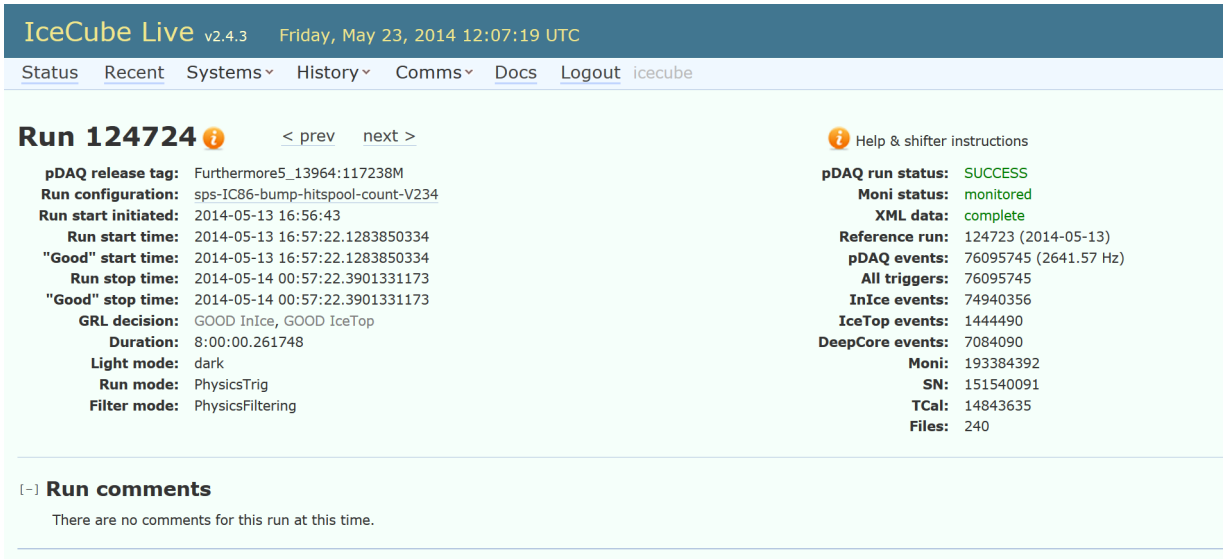


Figure A.1: Monitoring summary of a 8 h data-taking period (run).

The monitoring interface was implemented as a subsystem in IceCube's control software *IceCube Live* which is based on the Django web framework[1]. The monitoring backbone stores the information at the IceCube data storage in the form of XML-files[2] and ROOT-

[1] https://www.djangoproject.com/

[2] Extensible Markup Language: a syntax to store information in text files.

files[3]. The data is processed and stored in the MySQL database of IceCube Live, from where it is requested for display in the web interface. Figure A.1 shows the upper part of the page providing general information about a 8 h data-taking period (run).

Trigger rates

Show abnormal triggers only

Trigger type	Rate	Reference rate	Deviation
GLOBAL_MERGED	2107.36 ± 0.28	2105.59 ± 0.52	3.00 σ (0.08%)
GLOBAL_THROUGHPUT	8204.99 ± 0.55	8198.6 ± 1.0	5.50 σ (0.08%)
ICE_TOP_CALIBRATION	29.099 ± 0.033	28.961 ± 0.062	1.95 σ (0.48%)
ICE_TOP_MIN_BIAS	0.2622 ± 0.0032	0.2596 ± 0.006	0.38 σ (0.99%)
ICE_TOP_SIMPLE_MULTIPLICITY	23.96 ± 0.031	23.82 ± 0.058	2.15 σ (0.59%)
IN_ICE_SIMPLE_MULTIPLICITY	2064.5 ± 0.27	2062.77 ± 0.51	2.96 σ (0.08%)
IN_ICE_SIMPLE_MULTIPLICITY_DEEPCORE	250.785 ± 0.095	250.87 ± 0.18	-0.40 σ (-0.03%)
IN_ICE_SLOW_PARTICLE	12.017 ± 0.021	12.058 ± 0.039	-0.93 σ (-0.34%)
IN_ICE_STRING	2187.05 ± 0.28	2185.44 ± 0.53	2.69 σ (0.07%)
IN_ICE_UNBIASED	0.0033 ± 0.0003	0.0033 ± 0.0007	-
IN_ICE_VOLUME	3640.37 ± 0.36	3637.79 ± 0.68	3.33 σ (0.07%)

Filter rates

Show abnormal filters only

Filter type	Rate	Reference rate	Deviation
CASCADEFILTER_13	32.826 ± 0.035	32.684 ± 0.065	1.94 σ (0.43%)
DEEPCOREFILTER_13	16.465 ± 0.024	16.503 ± 0.046	-0.73 σ (-0.23%)
DEEPCOREFILTER_TWOLAYEREXP_13	35.881 ± 0.036	35.927 ± 0.068	-0.60 σ (-0.13%)
EHEFILTER_13	1.0028 ± 0.006	1.015 ± 0.011	-0.92 σ (-1.16%)
FILTERMINBIAS_13	2.6421 ± 0.0098	2.647 ± 0.018	-0.22 σ (-0.17%)
FIXEDRATEFILTER_13	0.0033 ± 0.0003	0.0033 ± 0.0007	-
FSSCANDIDATE_13	0.0 ± 0.0	-	-
FSSFILTER_13	184.109 ± 0.082	184.05 ± 0.15	0.35 σ (0.03%)
GCFILTER_13	92.212 ± 0.058	94.24 ± 0.11	-16.34 σ (-2.15%)
I3DAQDECODEEXCEPTION	0.0 ± 0.0	-	-
ICETOP_INFILL_STA3_13	0.254 ± 0.003	0.2512 ± 0.0057	0.43 σ (1.11%)
ICETOPSTA3_13	1.2664 ± 0.0068	1.246 ± 0.013	1.40 σ (1.60%)
ICETOPSTA5_13	3.073 ± 0.011	3.026 ± 0.02	2.10 σ (1.55%)
INICESMT_ICETOPCOINCIDENCE_13	0.8982 ± 0.0057	0.907 ± 0.011	-0.75 σ (-1.01%)
LID	0.0 ± 0.0	-	-
LOWUP_13	27.132 ± 0.031	27.092 ± 0.059	0.59 σ (0.15%)
MOONFILTER_13	53.405 ± 0.044	48.502 ± 0.079	54.28 σ (10.11%)
MUONFILTER_13	34.722 ± 0.036	34.777 ± 0.067	-0.73 σ (-0.16%)
OFUFILTER_14	0.0032 ± 0.0003	0.004 ± 0.0007	-1.05 σ (-20.00%)
ONLINEL2FILTER_14	5.448 ± 0.014	5.449 ± 0.026	-0.03 σ (-0.02%)
SDST_FILTERMINBIAS_13	456.11 ± 0.13	455.68 ± 0.24	1.57 σ (0.09%)
SDST_ICETOP_INFILL_STA3_13	2.4799 ± 0.0095	2.467 ± 0.018	0.65 σ (0.53%)
SDST_ICETOPSTA3_13	12.701 ± 0.022	12.534 ± 0.04	3.67 σ (1.33%)
SDST_INICESMT_ICETOPCOINCIDENCE_13	76.625 ± 0.053	75.349 ± 0.098	11.44 σ (1.69%)
SLOPFILTER_13	12.015 ± 0.021	12.058 ± 0.039	-0.95 σ (-0.35%)
SUNFILTER_13	0.0 ± 0.0	-	-
VEF_13	11.938 ± 0.021	12.027 ± 0.039	-2.01 σ (-0.74%)

Figure A.2: Monitoring trigger/filter rates: Large deviations from a reference run are shown in red.

Moreover, all trigger and filter rates are displayed as shown in Figure A.2. The simple multiplicity trigger and the muon filter, both touched on in this work, are given as *IN_ICE_SIMPLE_MULTIPLICITY* and *MUONFILTER_13*, respectively. The rates are compared to a reference run to reveal potential deviations, which are then shown in

[3]Readable with the ROOT data analysis framework [123].

red. By clicking a trigger/filter name, the rates are accessible as a function of time as well as a 40 day history.

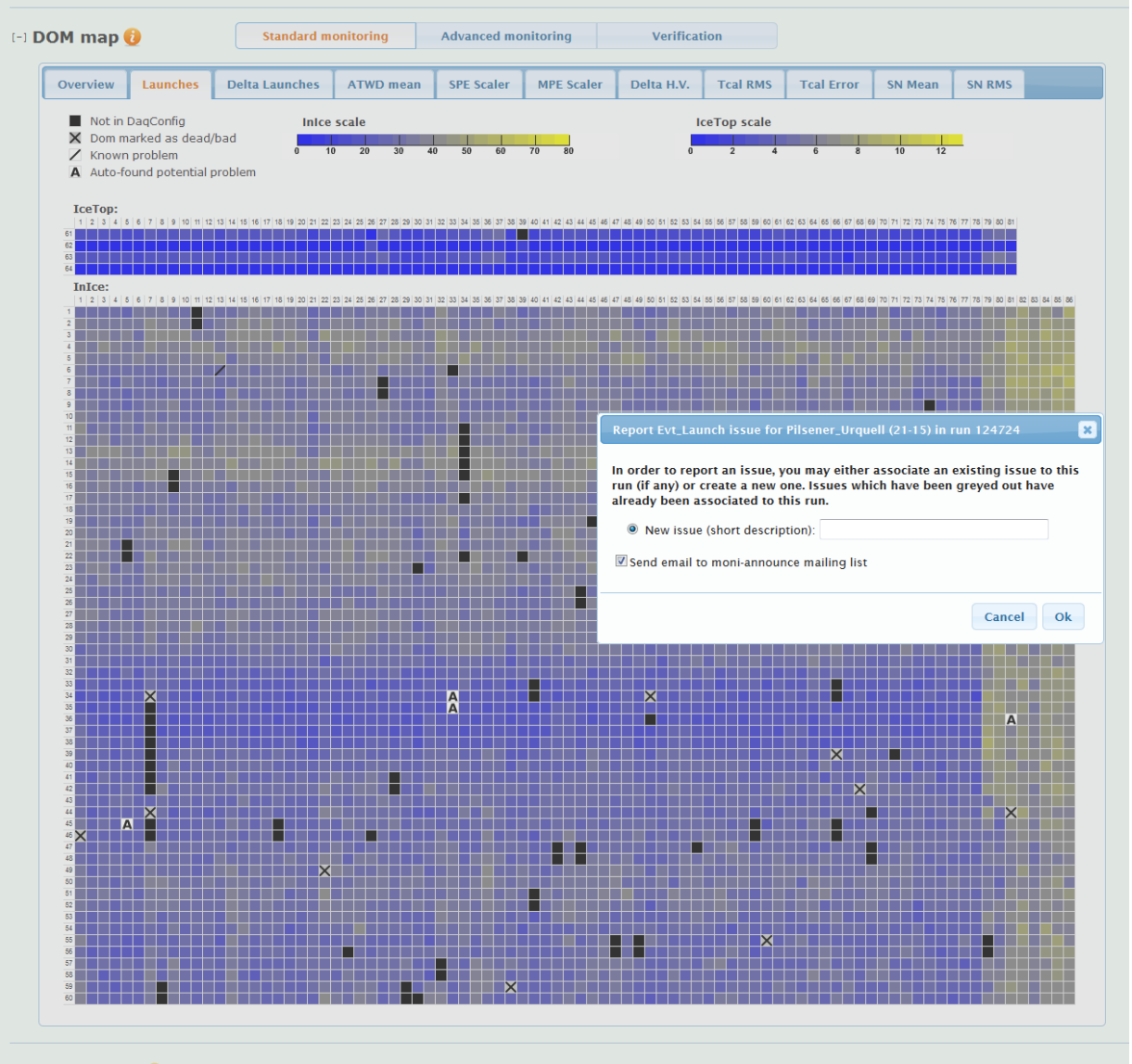
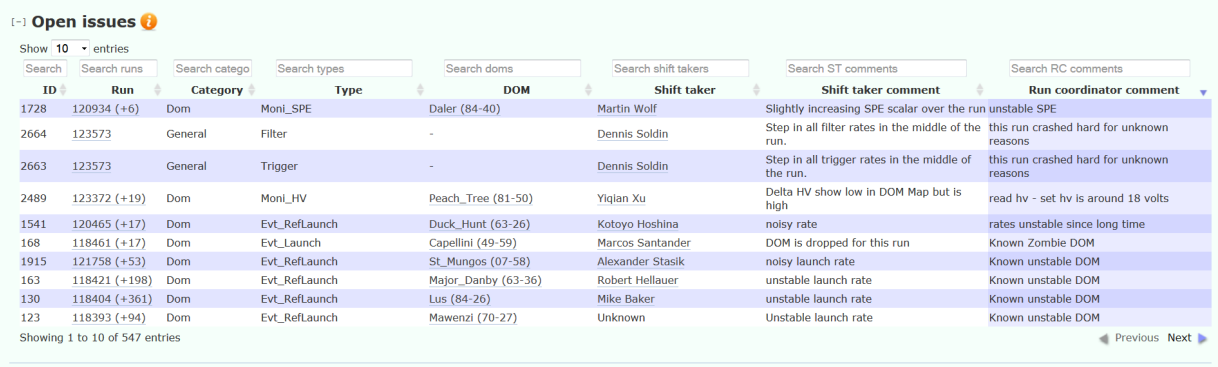


Figure A.3: Monitoring DOM map.

The DOM map in Figure A.3 shows the status and the rates of all DOMs in IceCube and IceTop. Deactivated DOMs are shown in black and DOMs with permanent serious problems are shown with a cross. They are not deactivated as they still provide coincidence information to their neighbors. An automatically detected potential problem in a run is marked with an A. The selected tab in the picture shows *DOM launches*, which is the rate at which a DOM participated in triggered events. Higher rates are obtained from the high-efficiency DeepCore DOMs on the right (strings 79 to 86) which are located below the big dust layer. The dust layer appears as a dark horizontal band in the standard IceCube DOMs. Single DOM rates as a function of time are accessible by clicking a specific DOM.

Figure A.3 additionally shows the dialog box for reporting potential issues in a DOM. A quick comment describing the problem can be given before it is stored in the database. Issues can be reported in the same way for trigger, filter and general information. A list displays these entries as shown in Figure A.4. The IceCube run coordinator can modify and delete issues as well as leave comments which also appear in the list.



[-] **Open issues**

Show 10 entries

Search	Search runs	Search catego	Search types	Search doms	Search shift takers	Search ST comments	Search RC comments
ID	**Run**	**Category**	**Type**	**DOM**	**Shift taker**	**Shift taker comment**	**Run coordinator comment**
1728	120934 (+6)	Dom	Moni_SPE	Daler (84-40)	Martin Wolf	Slightly increasing SPE scalar over the run	unstable SPE
2664	123573	General	Filter	-	Dennis Soldin	Step in all filter rates in the middle of the run.	this run crashed hard for unknown reasons
2663	123573	General	Trigger	-	Dennis Soldin	Step in all trigger rates in the middle of the run.	this run crashed hard for unknown reasons
2489	123372 (+19)	Dom	Moni_HV	Peach_Tree (81-50)	Yiqian Xu	Delta HV show low in DOM Map but is high	read hv - set hv is around 18 volts
1541	120465 (+17)	Dom	Evt_RefLaunch	Duck_Hunt (63-26)	Kotoyo Hoshina	noisy rate	rates unstable since long time
168	118461 (+17)	Dom	Evt_Launch	Capellini (49-59)	Marcos Santander	DOM is dropped for this run	Known Zombie DOM
1915	121758 (+53)	Dom	Evt_RefLaunch	St_Mungos (07-58)	Alexander Stasik	noisy launch rate	Known unstable DOM
163	118421 (+198)	Dom	Evt_RefLaunch	Major_Danby (63-36)	Robert Hellauer	unstable launch rate	Known unstable DOM
130	118404 (+361)	Dom	Evt_RefLaunch	Lus (84-26)	Mike Baker	unstable launch rate	Known unstable DOM
123	118393 (+94)	Dom	Evt_RefLaunch	Mawenzi (70-27)	Unknown	Unstable launch rate	Known unstable DOM

Showing 1 to 10 of 547 entries

Previous Next

Figure A.4: Reported data quality and detector issues.

IceCube monitoring is an collaboration wide effort and organized in a shift plan. The shift-taker in charge supervises the most recent data-taking periods on a daily base through the interactive interface developed in the course of this work. The reported issues are a valuable input for the classification of good runs and the identification of potential problems.

Appendix B

Additional tables and figures

Some tables and figures have been shifted to the appendix for better readability in the main sections. They provide complementary information and are referenced at the corresponding locations in the thesis.

Sample size (n)	**Critical values $K_{\alpha,n}$ for different significant levels α** 0.20	0.15	0.10	0.05	0.01
1	0.900	0.925	0.950	0.975	0.995
2	0.684	0.726	0.776	0.842	0.929
3	0.565	0.597	0.642	0.708	0.828
4	0.494	0.525	0.564	0.624	0.733
5	0.446	0.474	0.510	0.565	0.669
6	0.410	0.436	0.470	0.521	0.618
7	0.381	0.405	0.438	0.486	0.577
8	0.358	0.381	0.411	0.457	0.543
9	0.339	0.360	0.388	0.432	0.514
10	0.322	0.342	0.368	0.410	0.490
11	0.307	0.326	0.352	0.391	0.468
12	0.295	0.313	0.338	0.375	0.450
13	0.284	0.302	0.325	0.361	0.433
14	0.274	0.292	0.314	0.349	0.418
15	0.266	0.283	0.304	0.338	0.404
16	0.258	0.274	0.295	0.328	0.392
17	0.250	0.266	0.286	0.318	0.387
18	0.244	0.259	0.278	0.309	0.371
19	0.237	0.252	0.272	0.301	0.363
20	0.231	0.246	0.264	0.294	0.356

Table B.1: Critical values for the Kolmogorov-Smirnov test [153]. Referenced in Section 3.5.1.

N_{trees}	Maximal depth	Prune strength	Signal efficiency at bkg eff.(error): @B=0.01	@B=0.10	@B=0.30
800	5	-	0.489(03)	0.787(02)	0.911(02)
800	4	-	0.494(03)	0.787(02)	0.911(02)
400	5	-	0.486(03)	0.782(02)	0.910(02)
800	6	-	0.503(03)	0.784(02)	0.907(02)
400	4	-	0.487(03)	0.783(02)	0.909(02)
400	6	-	0.495(03)	0.781(02)	0.907(02)
800	3	-	0.480(03)	0.773(02)	0.913(01)
400	3	-	0.470(03)	0.766(02)	0.909(02)
800	3	5	0.473(03)	0.768(02)	0.908(02)
800	2	-	0.462(03)	0.754(03)	0.900(02)
400	3	5	0.451(03)	0.753(03)	0.900(02)
400	2	-	0.433(03)	0.735(03)	0.890(02)

Table B.2: Signal efficiency of different BDT configurations for the downgoing BDT2 at three different background efficiencies evaluated for the independent test sample. The BDTs from the original IC79 sample still use pruning noticeable by the given prune strength. Referenced in Section 4.3.3.

Component	Version or source path
IceRec meta-project	04-01-02
Shield, DeepCore_Filter and MuE	replaced from *svn/sandbox/gross/IC79_MUON_L3_META/releases/V01-00-02/*
MuEX	*svn/sandbox/schatto/mue/wreco/*
spline-reco	*svn/sandbox/schatto/spline-reco_IC86L3/spline-reco/*
small-tree	*svn/sandbox/Sirin/small-tree-IC79/branches/IC79_MUON_L3_v2*
splines	*/net/user/schatto/InfBareMu_mie_spline/*
L3b scripts	*svn/sandbox/schatto/IC79L3b/*

Table B.3: The software setup used for the reprocessing of the IC79 L3 data into L3b including the SplineMPE reconstruction. Referenced in Section 4.2.3.

cos(Zenith) range	BDT cut value	cos(Zenith) range	BDT cut value
0 to 0.0125	1.39	0.3375 to 0.35	1.182
0.0125 to 0.025	1.343	0.35 to 0.3625	1.173
0.025 to 0.0375	1.298	0.3625 to 0.375	1.178
0.0375 to 0.05	1.268	0.375 to 0.38	1.167
0.05 to 0.0625	1.262	0.38 to 0.3875	1.184
0.0625 to 0.075	1.253	0.3875 to 0.4	1.18
0.075 to 0.0875	1.249	0.4 to 0.4125	1.180
0.0875 to 0.093	1.254	0.4125 to 0.425	1.178
0.093 to 0.1	1.238	0.425 to 0.4375	1.179
0.1 to 0.11	1.242	0.4375 to 0.45	1.175
0.11 to 0.116	1.239	0.45 to 0.4625	1.179
0.116 to 0.125	1.24	0.4625 to 0.475	1.181
0.125 to 0.1305	1.231	0.475 to 0.4875	1.175
0.1305 to 0.1392	1.236	0.4875 to 0.5	1.177
0.1392 to 0.145	1.221	0.5 to 0.5125	1.177
0.145 to 0.1625	1.23	0.5125 to 0.525	1.177
0.1625 to 0.175	1.219	0.525 to 0.5375	1.175
0.175 to 0.1875	1.222	0.5375 to 0.55	1.175
0.1875 to 0.2	1.216	0.55 to 0.5625	1.175
0.2 to 0.2125	1.212	0.5625 to 0.575	1.175
0.2125 to 0.225	1.212	0.575 to 0.5875	1.175
0.225 to 0.2364	1.212	0.5875 to 0.6	1.175
0.2364 to 0.2419	1.202	0.6 to 0.6125	1.175
0.2419 to 0.25	1.206	0.6125 to 0.625	1.175
0.25 to 0.2588	1.212	0.625 to 0.6375	1.175
0.2588 to 0.264	1.220	0.6375 to 0.65	1.177
0.264 to 0.27	1.182	0.65 to 0.6625	1.177
0.27 to 0.275	1.186	0.6625 to 0.675	1.18
0.275 to 0.2875	1.20	0.675 to 0.6875	1.18
0.2875 to 0.3	1.191	0.6875 to 0.7	1.179
0.3 to 0.3125	1.19	0.7 to 0.7125	1.176
0.3125 to 0.325	1.19	≥ 0.7125	1.175
0.325 to 0.3375	1.189		

Table B.4: Adjusted BDT cut values in the downgoing region delivering a seamless transition at the horizon. Referenced in Section 4.3.6.

Zenith range	BDT cut value	
	hard spectrum BDT	soft spectrum BDT
89.5° to 130°	1.15	1.275
89° to 89.5°	1.155	1.28
88.5° to 89°	1.16	1.285
88° to 88.5°	1.17	1.295
87.5° to 88°	1.185	1.31
87° to 87.5°	1.2	1.325
86.5° to 87°	1.215	1.34
86° to 86.5°	1.245	1.37
85.5° to 86°	1.275	1.4
85° to 85.5°	1.335	1.46

Table B.5: BDT cut values for the horizontal upgoing BDTs in the overlap zone with the downgoing BDT. Referenced in Section 4.3.6.

Name	Ra [°]	Dec [°]	class	F_γ	α_γ	L_γ	N_ν
GB6 J0043+3426	10.9409	34.4394	-	1.9	2.12	1.2×10^{47}	1.07
PKS 0215+015	34.4833	1.73026	LSP	5.66	2.15	1.6×10^{48}	1.803
PKS 0454-234	74.268	-23.4271	-	22.7	2.21	1.4×10^{48}	0.6085
GB6 J0654+5042	103.649	50.7181	LSP	2.8	2.04	3.1×10^{47}	2.377
MG2 J071354+1934	108.512	19.5654	LSP	5.19	2.22	6.3×10^{46}	1.264
B2 0716+33	109.827	33.1079	LSP	8.29	2.06	2.2×10^{47}	7.645
4C +14.23	111.331	14.4429	LSP	9.63	2.18	6.3×10^{47}	3.398
PKS 0727-11	112.572	-11.6955	-	22	2.28	5.2×10^{48}	0.9398
PKS 0805-07	122.059	-7.84967	LSP	15	2.04	3.5×10^{48}	9.611
S4 0917+44	140.236	44.6968	LSP	9.1	2.27	5×10^{48}	1.124
4C +55.17	149.433	55.3817	LSP	11.2	1.95	4.3×10^{47}	19.81
TXS 1013+054	154.014	5.2291	-	4.4	2.08	1.1×10^{48}	3.087
4C +21.35	186.226	21.38	LSP	35.4	2.23	3×10^{47}	7.987
3C 273	187.277	2.04167	LSP	15.1	2.62	2.2×10^{46}	0.08222
PKS 1244-255	191.681	-25.7828	LSP	8.74	2.26	1.8×10^{47}	0.09651
3C 279	194.042	-5.79369	LSP	25.6	2.34	4.2×10^{47}	0.9096
OP 313	197.674	32.381	LSP	5.3	2.22	3.4×10^{47}	1.222
GB 1310+487	198.212	48.4722	LSP	8.99	2.11	9×10^{46}	4.333
PKS B1424-418	217.012	-42.1036	LSP	14.7	2.13	2.3×10^{48}	0.2636
PKS 1454-354	224.363	-35.6709	LSP	11.9	2.27	2×10^{48}	0.05042
PKS 1502+106	226.095	10.4883	LSP	40.1	2.21	1.2×10^{49}	9.804
PKS 1510-08	228.207	-9.10336	LSP	40.6	2.39	2.7×10^{47}	0.7181
B2 1520+31	230.542	31.7436	LSP	17.6	2.37	4.4×10^{48}	1.201
PKS 1622-253	246.428	-25.441	-	8.15	2.34	3.1×10^{47}	0.0366
4C +38.41	248.809	38.1712	LSP	11.6	2.41	5.3×10^{48}	0.5272
PMN J1802-3940	270.664	-39.6791	LSP	16.9	2.23	2.2×10^{48}	0.09716
PKS 1830-211	278.413	-21.0749	-	11.9	2.62	2.4×10^{49}	0.004969
S4 1849+67	282.35	67.1035	LSP	7.36	2.23	1.6×10^{47}	1.067
TXS 1920-211	290.884	-21.0869	LSP	8.33	2.24	3.8×10^{47}	0.1822
PKS 2023-07	306.416	-7.60522	LSP	10.1	2.32	1.8×10^{48}	0.3923
PKS 2052-47	314.068	-47.2549	LSP	8.63	2.35	2.1×10^{48}	0.007158
PKS 2201+171	330.869	17.4363	LSP	6.53	2.23	5×10^{47}	1.514
3C 454.3	343.497	16.1526	LSP	96.5	2.38	5.8×10^{48}	5.873

Table B.6: FSRQ catalog: The table lists object name, position in right ascension and declination (Ra and Dec, epoch J2000), gamma-flux F_γ from 1 GeV to 100 GeV in units of $10^{-9}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and the fitted power law index α_γ as stated by Fermi [55]. The luminosity L_γ from 0.1 GeV to 100 GeV in units of $\mathrm{erg}\,\mathrm{s}^{-1}$ is calculated according to Equation 5.9 and N_ν is defined in Equation 5.7. Referenced in Section 5.4.

Name	**Ra** [°]	**Dec** [°]	**class**	F_γ	α_γ	L_γ	N_ν
MG4 J000800+4712	1.97432	47.2298	LSP	2.13	2.1	6.4×10^{45}	1.119
PKS 0019+058	5.64299	6.12368	LSP	2.22	2.09	2.2×10^{36}	1.47
PKS 0118-272	20.1239	-27.0149	LSP	4.35	1.93	4.5×10^{36}	3.117
TXS 0141+268	26.1561	27.0793	LSP	3.4	2.17	3.5×10^{36}	1.304
AO 0235+164	39.6753	16.6241	LSP	18.7	2.12	9.4×10^{47}	10.8
PMN J0334-3725	53.5848	-37.4685	LSP	3.42	1.99	3.5×10^{36}	0.5455
PKS 0426-380	67.1709	-37.9334	LSP	31.1	2.04	2.1×10^{48}	2.41
MG2 J043337+2905	68.3886	29.0864	LSP	4.31	2.04	4.3×10^{36}	4.991
PKS 0537-441	84.706	-44.0838	LSP	37.1	2.09	1.6×10^{48}	1.043
GB6 J0712+5033	108.228	50.5366	LSP	1.73	2.06	1.7×10^{36}	1.244
PKS 0735+17	114.524	17.7026	LSP	5.16	2.05	4.1×10^{46}	6.195
S4 0814+42	124.573	42.3983	LSP	7.21	2.14	7.3×10^{36}	2.831
PKS 0829+046	127.987	4.48609	LSP	5	2.18	4.3×10^{45}	1.325
CRATES J0856-1105	134.173	-11.0957	LSP	2.89	2.05	2.9×10^{36}	1.386
4C +01.28	164.615	1.56646	LSP	5.12	2.22	2.8×10^{47}	0.8615
S4 1250+53	193.276	53.0467	LSP	3.73	1.97	3.8×10^{36}	5.333
PMN J1326-5256	201.7	-52.9149	LSP	3.88	2.23	4.1×10^{36}	0.01421
AP Librae	229.435	-24.36	LSP	5.16	2.05	3.4×10^{44}	0.8253
PKS 1519-273	230.676	-27.5229	LSP	3.55	2.22	5.1×10^{47}	0.05779
PMN J1603-4904	240.971	-49.0784	LSP	12.9	2.04	1.3×10^{37}	0.6984
PKS 1717+177	259.83	17.7425	LSP	2.23	1.84	1.4×10^{45}	20.42
OT 081	267.876	9.64001	LSP	4.56	2.24	1.6×10^{46}	0.7967
S5 1803+784	270.147	78.4834	LSP	4.45	2.23	1.3×10^{47}	0.5805
TXS 2036-109	309.786	-10.7789	LSP	2.22	1.99	2.2×10^{36}	2.217
B2 2234+28A	339.102	28.4755	LSP	5.08	2.22	1.7×10^{47}	1.282
PKS 2233-148	339.133	-14.5331	LSP	3.86	2.24	4.1×10^{36}	0.1906
B3 2322+396	351.331	39.959	LSP	2.33	1.99	2.4×10^{36}	3.715

Table B.7: LSP BL Lac catalog: The table lists object name, position in right ascension and declination (Ra and Dec, epoch J2000), gamma-flux F_γ from 1 GeV to 100 GeV in units of $10^{-9}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and the fitted power law index α_γ as stated by Fermi [55]. The luminosity L_γ from 0.1 GeV to 100 GeV in units of $\mathrm{erg}\,\mathrm{s}^{-1}$ is calculated according to Equation 5.9 and N_ν is defined in Equation 5.7. Referenced in Section 5.4.

Name	Ra [°]	Dec [°]	class	F_γ	α_γ	L_γ	N_ν
GB6 J0100+0745	15.0633	7.77599	-	2	1.71	2.4×10^{36}	77.63
S2 0109+22	18.0327	22.7637	ISP	7.13	2.07	1.9×10^{46}	7.109
B3 0133+388	24.143	39.0892	HSP	4.53	1.69	5.6×10^{36}	125.6
3C 66A	35.6617	43.0358	ISP	25.6	1.91	2.3×10^{37}	77.62
AO 0235+164	39.6753	16.6241	LSP	18.7	2.12	9.4×10^{47}	10.8
PKS 0301-243	45.8677	-24.1281	HSP	6.73	1.94	1.7×10^{46}	4.968
PKS 0426-380	67.1709	-37.9334	LSP	31.1	2.04	2.1×10^{48}	2.41
PKS 0447-439	72.371	-43.8384	HSP	11.4	1.86	1.7×10^{46}	8.986
1ES 0502+675	77.0131	67.6252	HSP	2.42	1.49	2.4×10^{46}	225.2
VER J0521+211	80.449	21.2204	ISP	9.04	1.93	9.3×10^{36}	33.54
PKS 0537-441	84.706	-44.0838	LSP	37.1	2.09	1.6×10^{48}	1.043
1ES 0647+250	102.699	25.0965	HSP	1.78	1.59	2.5×10^{36}	199.5
S5 0716+71	110.476	71.3496	ISP	18.3	2.08	1.7×10^{37}	8.729
PKS 0735+17	114.524	17.7026	LSP	5.16	2.05	4.1×10^{46}	6.195
S4 0814+42	124.573	42.3983	LSP	7.21	2.14	7.3×10^{36}	2.831
1H 1013+498	153.788	49.4324	HSP	7.8	1.85	9.9×10^{45}	36.69
4C +01.28	164.615	1.56646	LSP	5.12	2.22	2.8×10^{47}	0.8615
Mkn 421	166.12	38.2134	HSP	29.7	1.77	8.9×10^{44}	386.7
RBS 0958	169.305	20.2269	HSP	1.86	1.7	1.3×10^{45}	72.5
1ES 1215+303	184.467	30.1093	HSP	5.49	2.02	2.9×10^{45}	7.882
PG 1218+304	185.348	30.1787	HSP	2.81	1.71	3.7×10^{45}	82.88
W Comae	185.374	28.2391	ISP	5.54	2.02	1.8×10^{45}	8.201
Ton 116	190.78	36.4506	HSP	1.88	1.7	2.3×10^{36}	52.7
BZB J1418-0233	214.603	-2.5744	-	2.15	1.7	2.6×10^{36}	94.2
PKS 1424+240	216.76	23.795	ISP	11.5	1.78	1.3×10^{37}	188.1
AP Librae	229.435	-24.36	LSP	5.16	2.05	3.4×10^{44}	0.8253
GB6 J1542+6129	235.729	61.4887	ISP	6.37	1.97	6.5×10^{36}	8.29
PG 1553+113	238.942	11.1902	HSP	14	1.67	1.8×10^{37}	959.4
PMN J1603-4904	240.971	-49.0784	LSP	12.9	2.04	1.3×10^{37}	0.6984
Mkn 501	253.481	39.7631	HSP	8.77	1.74	3.2×10^{44}	151.5
1ES 1959+650	300.02	65.1572	HSP	5.88	1.94	3.7×10^{44}	9.807
MAGIC J2001+435	300.288	43.8791	ISP	11.8	1.9	1.2×10^{37}	38.34
MH 2136-428	324.841	-42.608	ISP	5.57	2.1	5.6×10^{36}	0.1654
PKS 2155-304	329.715	-30.2191	HSP	23.5	1.84	1×10^{46}	47.43
BL Lacertae	330.707	42.2675	ISP	10.5	2.26	1.3×10^{45}	1.497
B2 2234+28A	339.102	28.4755	LSP	5.08	2.22	1.7×10^{47}	1.282
RGB J2243+203	340.998	20.3562	HSP	3.93	1.75	4.6×10^{36}	89.77

Table B.8: Hard γ-spectrum BL Lac catalog: The table lists object name, position in right ascension and declination (Ra and Dec, epoch J2000), gamma-flux F_γ from 1 GeV to 100 GeV in units of $10^{-9}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and the fitted power law index α_γ as stated by Fermi [55]. The luminosity L_γ from 0.1 GeV to 100 GeV in units of $\mathrm{erg}\,\mathrm{s}^{-1}$ is calculated according to Equation 5.9 and N_ν is defined in Equation 5.7. Referenced in Section 5.4.

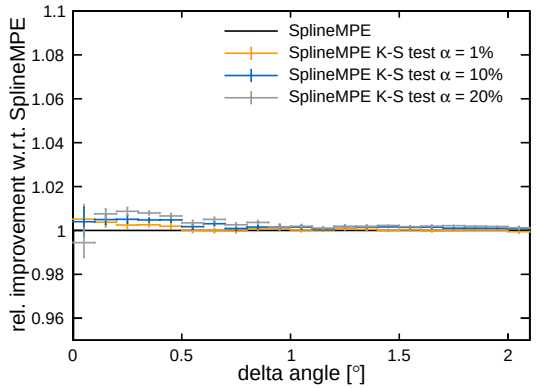


Figure B.1: Relative improvement of the cumulative point spread function with Kolmogorov-Smirnov cleaned pulses for an E^{-3} spectrum with respect to the default MPE reconstruction. Referenced in Section 3.5.1.

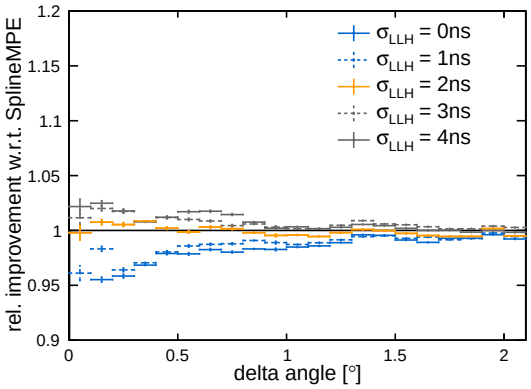


Figure B.2: Relative improvement of the cumulative point spread function with respect to the default MPE reconstruction with $\sigma_{\mathrm{PDF}} = 4\,\mathrm{ns}$ and $\sigma_{\mathrm{LLH}} = 0\,\mathrm{ns}$. Different σ_{LLH} values are shown for an E^{-3} neutrino spectrum, while σ_{PDF} is fixed to 2 ns. Referenced in Section 3.5.4.

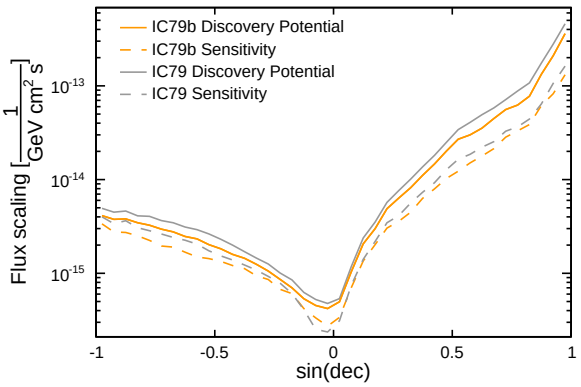


Figure B.3: E^{-1} sensitivity and discovery potential of the IC79b sample compared to the old IC79 sample. Referenced in Section 4.4.3.

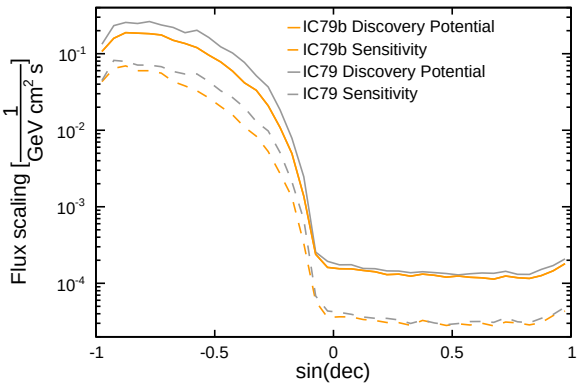


Figure B.4: E^{-3} sensitivity and discovery potential of the IC79b sample compared to the old IC79 sample. Referenced in Section 4.4.3.

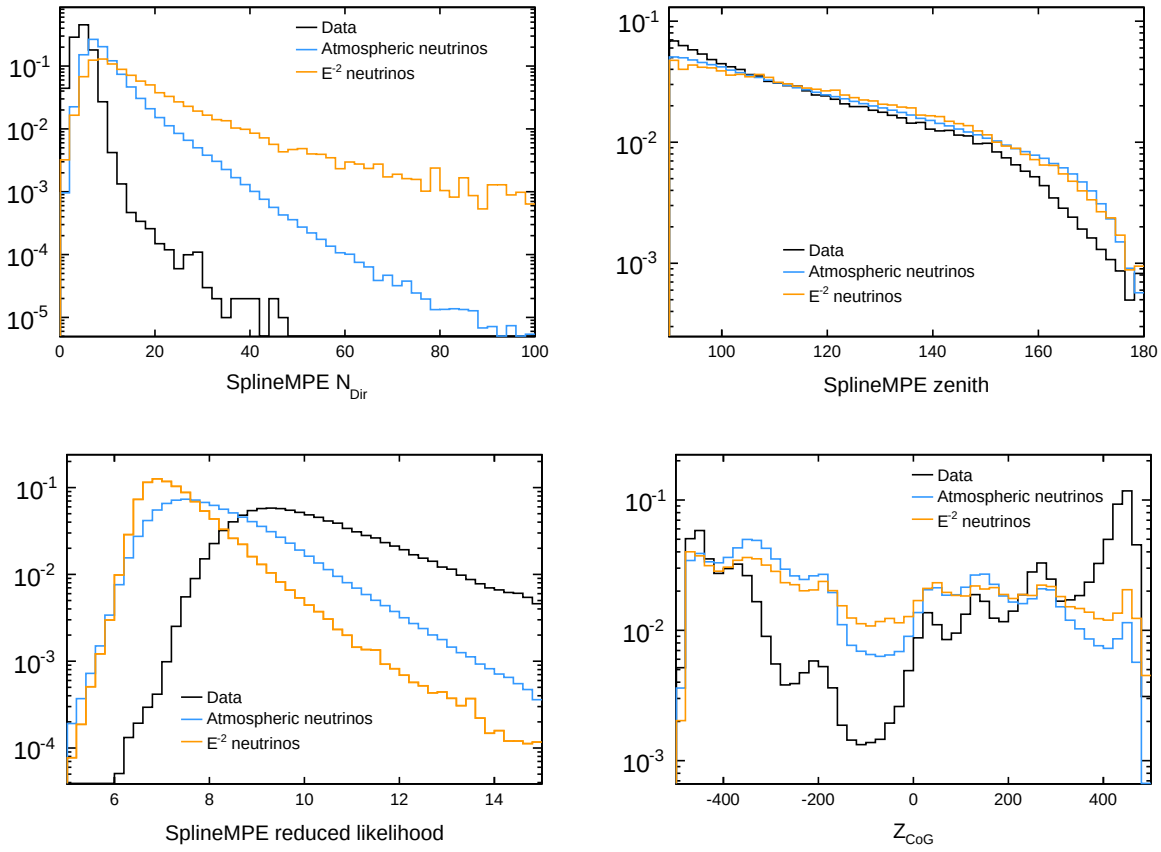


Figure B.5: Examples for signal-background separating variables on L3b: The atmospheric muon dominated IceCube data from the upgoing hemisphere on L3b is shown in black, the corresponding simulated atmospheric neutrino flux is shown in blue and an E^{-2} neutrino flux is shown in orange. The figure displays the normalized distributions of the variables N_{Dir} (upper left), SplineMPE zenith angle (upper right), SplineMPE $\mathcal{L}^{(2.5)}_{\mathrm{reduced}}$ (lower left) and Z_{CoG} (lower right) as specified in Section 4.1. Referenced in Section 4.3.2.

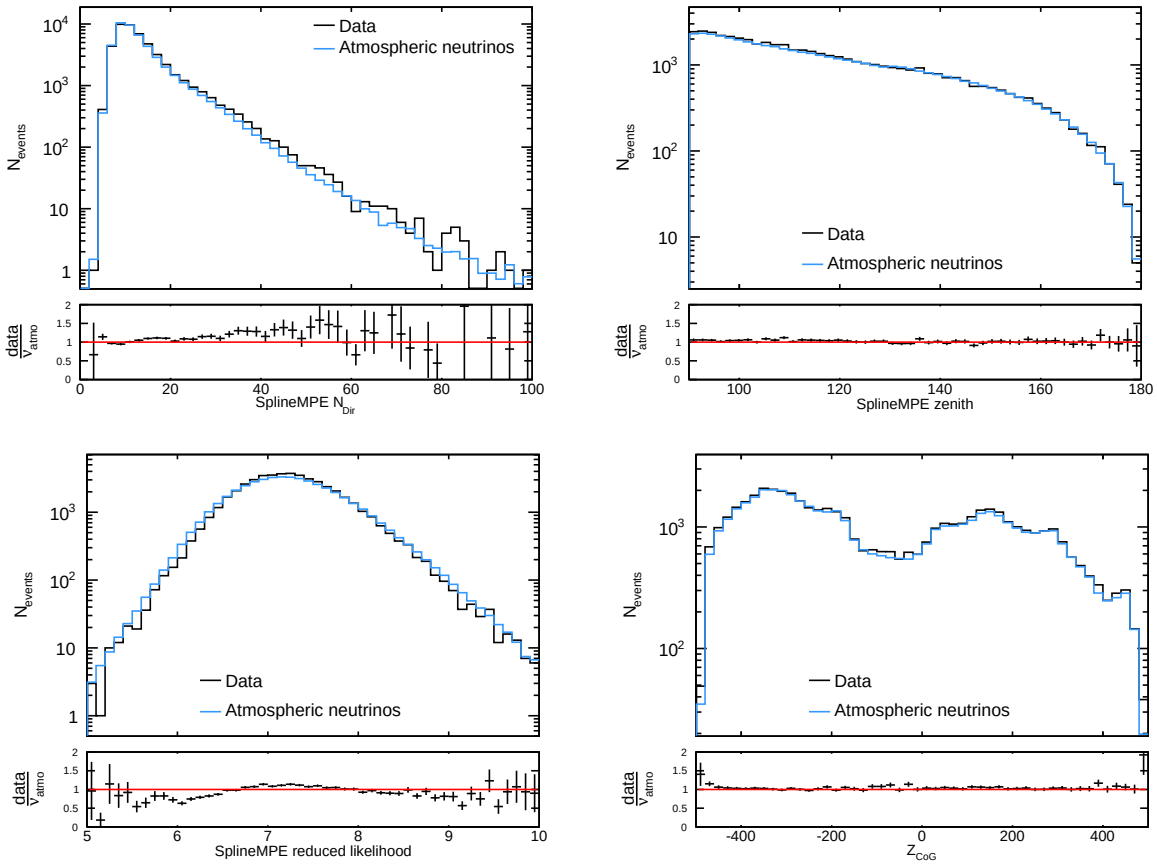


Figure B.6: Data-Monte Carlo comparison: The atmospheric neutrino dominated IC79b IceCube data from the upgoing hemisphere on the final selection level defined in Section 4.4 is shown in black while the corresponding simulated atmospheric neutrino flux is shown in blue. The figure shows the variables N_{Dir} (upper left), SplineMPE zenith angle (upper right), SplineMPE $\mathcal{L}^{(2.5)}_{\mathrm{reduced}}$ (lower left) and Z_{CoG} (lower right) as specified in Section 4.1. The deviation between data and simulation is displayed in the lower section of the plots. Referenced in Section 4.3.6.

References

[1] B. T. Cleveland, T. Daily, R. Davis, Jr., J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman. **Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector**. *ApJ*, **496**:505–526, March 1998. 4

[2] Q.R. Ahmad et al. **Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory**. *Phys.Rev.Lett.*, **87**:071301, 2001. 4

[3] E.W. Otten and C. Weinheimer. **Neutrino mass limit from tritium beta decay**. *Rept.Prog.Phys.*, **71**:086201, 2008. 4

[4] Bogdan Povh, Klaus Rith, Christoph Scholz, and Frank Zetsche. **Teilchen und Kerne**. *Physics and astronomy online library. Springer*, 2009. 4

[5] J. K. Becker. **High-energy neutrinos in the context of multimessenger astrophysics**. *Physics Reports*, **458**:173–246, March 2008. 4, 5, 8, 9, 14, 15

[6] B. Pontecorvo. **Neutrino Experiments and the Problem of Conservation of Leptonic Charge**. *Soviet Journal of Experimental and Theoretical Physics*, **26**:984, May 1968. 4

[7] Z. Maki, M. Nakagawa, and S. Sakata. **Remarks on the Unified Model of Elementary Particles**. *Progress of Theoretical Physics*, **28**:870–880, November 1962. 4

[8] Boris Kayser. **Neutrino mass, mixing, and flavor change**, 2002. 4

[9] John G. Learned and Sandip Pakvasa. **Detecting tau-neutrino oscillations at PeV energies**. *Astropart.Phys.*, **3**:267–274, 1995. 4

[10] G.L. Fogli, E. Lisi, A. Marrone, D. Montanino, A. Palazzo, et al. **Global analysis of neutrino masses, mixings and phases: entering the era of leptonic CP violation searches**. *Phys.Rev.*, **D86**:013012, 2012. 4

[11] Francesco Vissani, Giulia Pagliaroli, and Francesco L. Villante. **The fraction of muon tracks in cosmic neutrinos**. *JCAP*, **1309**:017, 2013. 4

[12] A. Yu. Smirnov. **The MSW effect and matter effects in neutrino oscillations**. *Phys.Scripta*, **T121**:57–64, 2005. 5

[13] R.N. Cahn, D.A. Dwyer, S.J. Freedman, W.C. Haxton, R.W. Kadel, et al. **White Paper: Measuring the Neutrino Mass Hierarchy**, 2013. 5

[14] V. F. Hess. **Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten**. *Physikalische Zeitschrift*, **13**:1084–1091, November 1912. 5

[15] Thomas K. Gaisser, Todor Stanev, and Serap Tilav. **Cosmic Ray Energy Spectrum from Measurements of Air Showers**. *Front.Phys.China.*, **8**:748–758, 2013. 5

[16] J. Beringer et al. **Review of Particle Physics (RPP)**. *Phys.Rev.*, **D86**:010001, 2012. 5, 6, 100

[17] D. J. Bird, S. C. Corbato, H. Y. Dai, J. W. Elbert, K. D. Green, M. A. Huang, D. B. Kieda, S. Ko, C. G. Larsen, E. C. Loh, M. Z. Luo, M. H. Salamon, J. D. Smith, P. Sokolsky, P. Sommers, J. K. K. Tang, and S. B. Thomas. **Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation**. *ApJ*, **441**:144–150, March 1995. 5

[18] B. Wiebel-Sooth, P. L. Biermann, and H. Meyer. **Cosmic rays. VII. Individual element spectra: prediction and data**. *Astronomy & Astrophysics*, **330**:389–398, February 1998. 5

[19] R. M. Kulsrud, J. P. Ostriker, and J. E. Gunn. **Acceleration of Cosmic Rays in Supernova Remnants**. *Physical Review Letters*, **28**:636–639, March 1972. 5

[20] T. Stanev. **Ultra High Energy Cosmic Rays:. Origin and Propagation**. In J. P. Wefel, T. Stanev, and M. M. Shapiro, editors, *Astrophysics at Ultra-High Energies*, pages 83–95, November 2007. 5, 8

[21] A. M. Hillas. **The Origin of Ultra-High-Energy Cosmic Rays**. *Annual Review of Astronomy & Astrophysics*, **22**:425–444, 1984. 5, 7, 8

[22] G. T. Zatsepin and V. A. Kuz'min. **Upper Limit of the Spectrum of Cosmic Rays**. *Soviet Journal of Experimental and Theoretical Physics Letters*, **4**:78, August 1966. 6

[23] K. Greisen. **End to the Cosmic-Ray Spectrum?** *Physical Review Letters*, **16**:748–750, April 1966. 6

[24] R. U. Abbasi, T. Abu-Zayyad, M. Allen, J. F. Amman, G. Archbold, K. Belov, J. W. Belz, S. Y. Ben Zvi, D. R. Bergman, S. A. Blake, O. A. Brusova, G. W. Burt, C. Cannon, Z. Cao, B. C. Connolly, W. Deng, Y. Fedorova, C. B. Finley, R. C. Gray, W. F. Hanlon, C. M. Hoffman, M. H. Holzscheiter, G. Hughes, P. Hüntemeyer, B. F. Jones, C. C. H. Jui, K. Kim, M. A. Kirn, E. C. Loh, M. M. Maestas, N. Manago, L. J. Marek, K. Martens, J. A. J. Matthews, J. N. Matthews, S. A. Moore, A. O'Neill, C. A. Painter, L. Perera, K. Reil, R. Riehle, M. Roberts, D. Rodriguez, N. Sasaki, S. R. Schnetzer, L. M. Scott, G. Sinnis, J. D. Smith, P. Sokolsky, C. Song, R. W. Springer, B. T. Stokes, S. B. Thomas, J. R. Thomas, G. B. Thomson, D. Tupa, S. Westerhoff, L. R. Wiencke, X. Zhang, and A. Zech. **First Observation of the Greisen-Zatsepin-Kuzmin Suppression**. *Physical Review Letters*, **100**(10):101101, March 2008. 6

[25] J. Abraham, P. Abreu, M. Aglietta, C. Aguirre, D. Allard, I. Allekotte, J. Allen, P. Allison, J. Alvarez-Muñiz, M. Ambrosio, and et al. **Observation of the Suppression of the Flux of Cosmic Rays above 4×10^{19} eV**. *Physical Review Letters*, **101**(6):061101, August 2008. 6

[26] ENRICO Fermi. **On the Origin of the Cosmic Radiation**. *Phys. Rev.*, **75**:1169–1174, Apr 1949. 6

[27] M. S. Longair. **High energy astrophysics. Volume 2. Stars, the Galaxy and the interstellar medium.** *Cambridge University Press*, Cambridge, 1994. 6, 7

[28] A. R. Bell. **The acceleration of cosmic rays in shock fronts. I**. *MNRAS*, **182**:147–156, January 1978. 6

[29] F. Halzen and E. Zas. **Neutrino fluxes from active galaxies: A Model independent estimate**. *Astrophys.J.*, **488**:669–674, 1997. 7, 12, 93

[30] Pierre Auger Collaboration, J. Abraham, P. Abreu, M. Aglietta, C. Aguirre, D. Allard, I. Allekotte, J. Allen, P. Allison, C. Alvarez, and et al. **Correlation of the Highest-Energy Cosmic Rays with Nearby Extragalactic Objects**. *Science*, **318**:938–, November 2007. 7

[31] M. Ackermann et al. **Detection of the Characteristic Pion-Decay Signature in Supernova Remnants**. *Science*, **339**:807, 2013. 8

[32] M. A. Markov and I. M. Zheleznykh. **On high energy neutrino physics in cosmic rays**. *Nuclear Physics A.*, **27**:385–394, October 1961. 8

[33] L. Köpke and IceCube Collaboration. **Supernova Neutrino Detection with IceCube**. *Journal of Physics Conference Series*, **309**(1):012029, August 2011. 9

REFERENCES

[34] K. S. Hirata, T. Kajita, M. Koshiba, M. Nakahata, Y. Oyama, N. Sato, A. Suzuki, M. Takita, Y. Totsuka, T. Kifune, T. Suda, K. Takahashi, T. Tanimori, K. Miyano, M. Yamada, E. W. Beier, L. R. Feldscher, W. Frati, S. B. Kim, A. K. Mann, F. M. Newcomer, R. Van Berg, W. Zhang, and B. G. Cortez. **Observation in the Kamiokande-II detector of the neutrino burst from supernova SN1987A**. *Phys. Rev. D*, **38**:448–458, Jul 1988. 9

[35] Eli Waxman and Abraham Loeb. **TeV neutrinos and GeV photons from shock breakout in supernovae**. *Phys.Rev.Lett.*, **87**:071101, 2001. 9

[36] K. Murase, T. A. Thompson, B. C. Lacki, and J. F. Beacom. **New class of high-energy transients from crashes of supernova ejecta with massive circumstellar material shells**. *Phys. Rev. D*, **84**(4):043003, August 2011. 9

[37] F. Aharonian et al. **Discovery of very-high-energy gamma-rays from the galactic centre ridge**. *Nature*, **439**:695–698, 2006. 9

[38] A.A. Abdo, Branden T. Allen, D. Berley, E. Blaufuss, S. Casanova, et al. **Discovery of TeV Gamma-Ray Emission from the Cygnus Region of the Galaxy**. *Astrophys.J.*, **658**:L33–L36, 2007. 9

[39] Francis Halzen and Aongus O Murchadha. **Neutrinos from Cosmic Ray Accelerators in the Cygnus Region of the Galaxy**. *Phys.Rev.*, **D76**:123003, 2007. 9

[40] Kevin Hurley, S.E. Boggs, D.M. Smith, R.C. Duncan, R. Lin, et al. **A Tremendous flare from SGR1806-20 with implications for short-duration gamma-ray bursts**. *Nature*, **434**:1098–1103, 2005. 9

[41] Francis Halzen, Hagar Landsman, and Teresa Montaruli. **TeV photons and neutrinos from giant soft-gamma repeaters flares**, 2005. 9

[42] W. Bednarek. **TeV Neutrinos from Microquasars in Compact Massive Binaries**. *ApJ*, **631**:466–470, September 2005. 9

[43] R. W. Klebesadel, I. B. Strong, and R. A. Olson. **Observations of Gamma-Ray Bursts of Cosmic Origin**. *ApJ*, **182**:L85, June 1973. 10

[44] C. Kouveliotou, C. A. Meegan, G. J. Fishman, N. P. Bhat, M. S. Briggs, T. M. Koshut, W. S. Paciesas, and G. N. Pendleton. **Identification of two classes of gamma-ray bursts**. *ApJ*, **413**:L101–L104, August 1993. 10

[45] D. Eichler, M. Livio, T. Piran, and D. N. Schramm. **Nucleosynthesis, neutrino bursts and gamma-rays from coalescing neutron stars**. *Nature*, **340**:126–128, July 1989. 10

[46] A. MacFadyen and S.E. Woosley. **Collapsars: Gamma-ray bursts and explosions in 'failed supernovae'**. *Astrophys.J.*, **524**:262, 1999. 10

[47] Bing Zhang. **Gamma-Ray Bursts in the Swift Era**. *Chin.J.Astron.Astrophys.*, **7**:1–50, 2007. 10

[48] R. Salvaterra, M. Della Valle, S. Campana, G. Chincarini, S. Covino, P. D'Avanzo, A. Fernández-Soto, C. Guidorzi, F. Mannucci, R. Margutti, C. C. Thöne, L. A. Antonelli, S. D. Barthelmy, M. de Pasquale, V. D'Elia, F. Fiore, D. Fugazza, L. K. Hunt, E. Maiorano, S. Marinoni, F. E. Marshall, E. Molinari, J. Nousek, E. Pian, J. L. Racusin, L. Stella, L. Amati, G. Andreuzzi, G. Cusumano, E. E. Fenimore, P. Ferrero, P. Giommi, D. Guetta, S. T. Holland, K. Hurley, G. L. Israel, J. Mao, C. B. Markwardt, N. Masetti, C. Pagani, E. Palazzi, D. M. Palmer, S. Piranomonte, G. Tagliaferri, and V. Testa. **GRB090423 at a redshift of z˜8.1**. *Nature*, **461**:1258–1260, October 2009. 10

[49] A. Franckowiak. **Searching for High-energy Neutrinos from Supernovae with IceCube and an Optical Follow-up Program**. PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät der Rheinischen Friedrich-Wilhelms-Universität, Bonn, Germany, 2011. 10, 17, 22

[50] Eli Waxman and John N. Bahcall. **High-energy neutrinos from cosmological gamma-ray burst fireballs**. *Phys.Rev.Lett.*, **78**:2292–2295, 1997. 10

[51] R. Abbasi, Y. Abdou, T. Abu-Zayyad, M. Ackermann, J. Adams, J. A. Aguilar, M. Ahlers, D. Altmann, K. Andeen, J. Auffenberg, and et al. **An absence of neutrinos associated with cosmic-ray acceleration in γ-ray bursts**. *Nature*, **484**:351–354, April 2012. 10, 32

[52] Markus Boettcher. **Modeling the Spectral Energy Distributions and Variability of Blazars**. 2012 Fermi & Jansky Proceedings - eConf C1111101, 2012. 10, 13, 14, 15, 94

[53] C. K. Seyfert. **Nuclear Emission in Spiral Nebulae.** *ApJ*, **97**:28, January 1943. 10, 11

[54] C. Megan Urry and Paolo Padovani. **Unified schemes for radio-loud active galactic nuclei**. *Publ.Astron.Soc.Pac.*, **107**:803, 1995. 10, 11, 12

[55] The Fermi LAT Collaboration. **The Second Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope**. *ApJ*, **743**:171, December 2011. 10, 12, 94, 99, 100, 120, 121, 122

[56] Daniel J. Mortlock, Stephen J. Warren, Bram P. Venemans, Mitesh Patel, Paul C. Hewett, et al. **A luminous quasar at a redshift of z = 7.085**. *Nature*, **474**:616, 2011. 10

[57] K. I. Kellermann, R. Sramek, M. Schmidt, D. B. Shaffer, and R. Green. **VLA observations of objects in the Palomar Bright Quasar Survey**. *AJ*, **98**:1195–1207, October 1989. 10

[58] Diego F. Torres. **Extragalactic gamma-ray sources**, 2003. 12

[59] B. L. Fanaroff and J. M. Riley. **The morphology of extragalactic radio sources of high and low luminosity**. *Monthly Notices of the Royal Astronomical Society*, **167**:31P–36P, May 1974. 11

[60] M. S. Shaw, R. W. Romani, S. E. Healey, G. Cotter, P. F. Michelson, and A. C. S. Readhead. **Optical Spectroscopy of Bright Fermi LAT Blazars**. *ApJ*, **704**:477–484, October 2009. 13

[61] M. Stickel, P. Padovani, C. M. Urry, J. W. Fried, and H. Kuehr. **The complete sample of 1 Jansky BL Lacertae objects. I - Summary properties**. *ApJ*, **374**:431–439, June 1991. 11

[62] P. Giommi, P. Padovani, G. Polenta, S. Turriziani, V. D'Elia, and S. Piranomonte. **A simplified view of blazars: clearing the fog around long-standing selection effects**. *MNRAS*, **420**:2899–2911, March 2012. 11

[63] R. D. Blandford and R. L. Znajek. **Electromagnetic extraction of energy from Kerr black holes**. *MNRAS*, **179**:433–456, May 1977. 12

[64] Armen Atoyan and Charles D. Dermer. **High-energy neutrinos from photomeson processes in blazars**. *Phys.Rev.Lett.*, **87**:221102, 2001. 12, 14, 15, 94, 95, 106, 107, 108

[65] A. M. Hillas. **Evolution of ground-based gamma-ray astronomy from the early days to the Cherenkov Telescope Arrays**. *Astroparticle Physics*, **43**:19–43, March 2013. 12

[66] Diego F. Torres and Luis A. Anchordoqui. **Astrophysical origins of ultrahigh energy cosmic rays**. *Rept.Prog.Phys.*, **67**:1663–1730, 2004. 12

[67] Todor Stanev. **Ultrahigh Energy Cosmic Rays: Origin and propagation**. *Proceedings of the 30th International Cosmic Ray Conference, Merida, Mexico*, 2007. 12

[68] A. Muecke, R.J. Protheroe, R. Engel, J.P. Rachen, and T. Stanev. **BL Lac Objects in the synchrotron proton blazar model**. *Astropart.Phys.*, **18**:593–613, 2003. 13, 14, 15, 94, 95, 106, 107, 108

[69] A.A. Abdo, M. Ackermann, M. Ajello, M. Axelsson, L. Baldini, et al. **The Spectral Energy Distribution of Fermi bright blazars**. *Astrophys.J.*, **716**:30–70, 2010. 13

[70] The Fermi-: A.A. Abdo et al., : V.A. Acciari et al., the GASP-WEBT consortium, and multi-wavelength partners. **Multi-wavelength Observations of the Flaring Gamma-ray Blazar 3C 66A in 2008 October**. *Astrophys.J.*, **726**:43, 2011. 14

[71] Markus Boettcher. **Modeling the Emission Processes in Blazars**. *Astrophys.Space Sci.*, **309**:95–104, 2007. 13

[72] Matthias Weidinger. **Systematic Modeling of Active Galactic Nuclei**. In *Systematic Modeling of Active Galactic Nuclei*. 33rd ICRC Track GA-TH (Rio de Janeiro), 2013. 13

[73] Sarira Sahu, Andres Felipe Osorio Oliveros, and Juan Carlos Sanabria. **Hadronic-origin orphan TeV flare from 1ES 1959+650**. *Phys.Rev.*, **D87**(10):103015, 2013. 14

[74] F.A. Aharonian. **TeV gamma rays from {BL} Lac objects due to synchrotron radiation of extremely high energy protons**. *New Astronomy*, **5**(7):377 – 395, 2000. 14

[75] A. Neronov and M. Ribordy. **IceCube sensitivity for neutrino flux from Fermi blazars in quiescent states**. *Phys. Rev. D*, **80**(8):083008, October 2009. 14, 15, 93, 94, 95, 101, 106, 107, 108

[76] F. Bernard. **Caracterisation Des Performances D'un Telescope Sous-marin A Neutrinos Pour La Detection De Cascades Contenues Dans Le Cadre Du Projet Antares**. PhD thesis, Université de la Méditerranée Aix-Marseille II, Marseille, France, 2000. 15

[77] J. Ahrens et al. **Muon track reconstruction and data selection techniques in AMANDA**. *Nucl.Instrum.Meth.*, **A524**:169–194, 2004. 15, 36, 41

[78] J. Abdallah et al. **A Precise measurement of the tau lifetime**. *Eur.Phys.J.*, **C36**:283–296, 2004. 15

[79] J. G. Learned and K. Mannheim. **High-energy neutrino astrophysics**. *Annual Review of Nuclear and Particle Science*, **50**:679–749, 2000. 16

[80] Raj Gandhi, Chris Quigg, Mary Hall Reno, and Ina Sarcevic. **Neutrino interactions at ultrahigh-energies**. *Phys.Rev.*, **D58**:093009, 1998. 16, 29

[81] The IceCube Collaboration. **An improved method for measuring muon energy using the truncated mean of dE/dx**. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **703**(0):190 – 198, 2013. 17, 18, 66, 69

[82] S. Iyer Dutta, M.H. Reno, I. Sarcevic, and D. Seckel. **Propagation of muons and taus at high-energies**. *Phys.Rev.*, **D63**:094020, 2001. 17, 18

[83] Paolo Lipari and Todor Stanev. **Propagation of multi-TeV muons**. *Phys. Rev. D*, **44**:3543–3554, Dec 1991. 17

[84] P. A. Čerenkov. **Visible Radiation Produced by Electrons Moving in a Medium with Velocities Exceeding that of Light**. *Phys. Rev.*, **52**:378–379, Aug 1937. 18

[85] C. A. Mead. **Quantum Theory of the Refractive Index**. *Phys. Rev.*, **110**:359–369, Apr 1958. 19

[86] Christopher Henrik V. Wiebusch. **The Detection of faint light in deep underwater neutrino telescopes**. PhD thesis, RWTH Aachen, Aachen, Germany, 1995. 19, 29, 55

[87] Johan Lundberg, P. Miocinovic, T. Burgess, J. Adams, S. Hundertmark, et al. **Light tracking for glaciers and oceans: Scattering and absorption in heterogeneous media with Photonics**. *Nucl.Instrum.Meth.*, **A581**:619–631, 2007. 19, 27, 30, 31, 41, 42, 43

[88] The IceCube Collaboration. **Optical properties of deep glacial ice at the South Pole**. *Journal of Geophysical Research: Atmospheres*, **111**(D13), 2006. 19, 24, 27, 28, 67

[89] Gustav Mie. **Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen**. *Annalen der Physik*, **330**(3):377–445, 1908. 19

[90] L. G. Henyey and J. L. Greenstein. **Diffuse radiation in the Galaxy**. *ApJ*, **93**:70–83, January 1941. 19

[91] P Liu. **A new phase function approximating to Mie scattering for radiative transport equations.** *Phys Med Biol*, **39**(6):1025–36, 1994. 19

[92] M.G. Aartsen et al. **Measurement of South Pole ice transparency with the IceCube LED calibration system**. *Nucl.Instrum.Meth.*, **A711**:73–89, 2013. 19, 28

[93] R. Abbasi et al. **Measurement of the atmospheric neutrino energy spectrum from 100 GeV to 400 TeV with IceCube**. *Phys.Rev.*, **D83**:012001, 2011. 20, 21

[94] Natalie Milke. **Unfolding of the atmospheric neutrino flux spectrum with the new program TRUEE and IceCube**. PhD thesis, Technische Universität Dortmund, Fachbereich Physik, Dortmund, Germany, 2012. 20, 21

[95] Patrick Berghaus. **Direct Measurement of the Atmospheric Muon Energy Spectrum with IceCube**. In *Direct Measurement of the Atmospheric Muon Energy Spectrum with IceCube*. Presented at the 31st ICRC, Lodz, Poland, July 2009, 2009. 20, 21

[96] Thomas K. Gaisser. **Atmospheric neutrino fluxes**. *Nucl.Phys.Proc.Suppl.*, **118**:109–117, 2003. 20, 21

[97] Thomas K. Gaisser. **Atmospheric Neutrinos**. *AIP Conf.Proc.*, **944**:140–142, 2007. 20

[98] Rikard Enberg, Mary Hall Reno, and Ina Sarcevic. **Prompt neutrino fluxes from atmospheric charm**. *Phys.Rev.*, **D78**:043005, 2008. 20, 21

[99] S.I. Sinegovsky, A.A. Kochanov, T.S. Sinegovskaya, A. Misaki, and N. Takahashi. **Atmospheric muon flux at PeV energies**. *Int.J.Mod.Phys.*, **A25**:3733–3740, 2010. 21

[100] Morihiro Honda, T. Kajita, K. Kasahara, S. Midorikawa, and T. Sanuki. **Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data**. *Phys.Rev.*, **D75**:043006, 2007. 21

[101] G.D. Barr, T.K. Gaisser, P. Lipari, Simon Robbins, and T. Stanev. **A Three - dimensional calculation of atmospheric neutrinos**. *Phys.Rev.*, **D70**:023006, 2004. 21

[102] E. V. Bugaev, V. A. Naumov, S. I. Sinegovskii, and E. S. Zaslavskaia. **Prompt leptons in cosmic rays**. *Nuovo Cimento C Geophysics Space Physics C*, **12**:41–73, February 1989. 21

[103] S. R. Klein. **IceCube: A Cubic Kilometer Radiation Detector**. *IEEE Transactions on Nuclear Science*, **56**:1141–1147, June 2009. 22, 23, 24, 25, 26, 59

[104] R. Abbasi, Y. Abdou, T. Abu-Zayyad, M. Ackermann, J. Adams, J. A. Aguilar, M. Ahlers, M. M. Allen, D. Altmann, K. Andeen, and et al. **The design and performance of IceCube DeepCore**. *Astroparticle Physics*, **35**:615–624, May 2012. 23

[105] K. Hanson and O. Tarasova. **Design and production of the IceCube digital optical module**. *Nucl.Instrum.Meth.*, **A567**:214–217, 2006. 23, 24

[106] R. Abbasi, Y. Abdou, T. Abu-Zayyad, J. Adams, J. A. Aguilar, M. Ahlers, K. Andeen, J. Auffenberg, X. Bai, M. Baker, and et al. **Calibration and characterization of the IceCube photomultiplier tube**. *Nuclear Instruments and Methods in Physics Research A*, **618**:139–152, June 2010. 24, 25, 58, 59

[107] John Kelley and Jim Braun. **Dom-Cal User Guide**, April 2005. 25

[108] R. Abbasi et al. **The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping**. *Nucl.Instrum.Meth.*, **A601**:294–316, 2009. 25, 58, 59

REFERENCES

[109] The IceCube Collaboration. **Search for Time-independent Neutrino Emission from Astrophysical Sources with 3 yr of IceCube Data**. *The Astrophysical Journal*, **779**(2):132, 2013. 26, 65, 70, 71, 75, 76, 78, 96, 97, 101, 104

[110] Nathan Whitehorn. **A Search for High-Energy Neutrino Emission from Gamma-Ray Bursts**. PhD thesis, University of Wisconsin–Madison, Wisconsin, Madison, 2012. 26, 27, 46

[111] N. E. Bramall, R. C. Bay, K. Woschnagg, R. A. Rohde, and P. B. Price. **A deep high-resolution optical log of dust, ash, and stratigraphy in South Pole glacial ice**. *Geophysical Research Letters*, **32**(21):n/a–n/a, 2005. 27

[112] P.B. Price. **Kinetics of conversion of air bubbles to air hydrate crystals in Antarctic ice**. *Science*, **267**:1802, 1995. 27

[113] Askhat Gazizov and Marek P. Kowalski. **ANIS: High energy neutrino generator for neutrino telescopes**. *Comput.Phys.Commun.*, **172**:203–213, 2005. 29

[114] Amanda Cooper-Sarkar, Philipp Mertsch, and Subir Sarkar. **The high energy neutrino cross-section in the Standard Model and its uncertainty**. *JHEP*, **1108**:042, 2011. 29

[115] D. Heck. **Extensive air shower simulations with CORSIKA and the influence of highenergy hadronic interaction models**, 2000. 29

[116] Dmitry Chirkin and Wolfgang Rhode. **Muon Monte Carlo: A High-precision tool for muon propagation through matter**, 2004. 29

[117] Dmitry Chirkin. **Photon tracking with GPUs in IceCube**. *Nucl.Instrum.Meth.*, **A725**:141–143, 2013. 30

[118] M.G. Aartsen et al. **Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector**. *Science*, **342**(6161):1242856, 2013. 31, 32

[119] M.G. Aartsen et al. **Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data**, 2014. 32

[120] The IceCube Collaboration. **Searches for Extended and Point-like Neutrino Sources with Four Years of IceCube Data**. *in preperation*, 2014. 32, 50, 63, 101, 104, 110

[121] R. Abbasi, Y. Abdou, T. Abu-Zayyad, J. Adams, J. A. Aguilar, M. Ahlers, K. Andeen, J. Auffenberg, X. Bai, M. Baker, and et al. **Time-dependent Searches for Point Sources of Neutrinos with the 40-string and 22-string Configurations of IceCube**. *ApJ*, **744**:1, January 2012. 32

[122] R. Abbasi et al. **Time-Integrated Searches for Point-like Sources of Neutrinos with the 40-String IceCube Detector**. *Astrophys.J.*, **732**:18, 2011. 32, 98, 101, 104

[123] Rene Brun and Fons Rademakers. **ROOT - An Object Oriented Data Analysis Framework**. In *ROOT - An Object Oriented Data Analysis Framework*. Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86, 1996. 37, 76, 78, 112

[124] Michael Soiron. **Untersuchungen zur Verbesserung der Myonenspur-Rekonstruktion in IceCube**. Master's thesis, Rheinisch-Westfälische Technische Hochschule Aachen, Mathematisch-Naturwissenschaftliche Fakultät, Aachen, Germany, 2012. 38

[125] D. Pandel. **Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal-Neutrinoteleskop NT-72**. Master's thesis, Humboldt-Universität zu Berlin, Berlin, Germany, 1996. 39

[126] B Efron. **Bootstrap methods: Another look at the Jackknife**. *Annals of Statistics*, **7**:1–26, 1979. 40

[127] R. C. Bay, R. A. Rohde, P. B. Price, and N. E. Bramall. **South Pole paleowind from automated synthesis of ice core records**. *Journal of Geophysical Research: Atmospheres*, **115**(D14), 2010. 42

[128] N. Whitehorn, J. van Santen, and S. Lafebre. **Penalized splines for smooth representation of high-dimensional Monte Carlo datasets**. *Computer Physics Communications*, **184**:2214–2220, September 2013. 46, 47

[129] W.T. Eadie and F. James. **Statistical Methods in Experimental Physics**. *World Scientific Publishing Company Incorporated*, 2006. 53

[130] Joschua Hellemeier. **Effect of noise on IceCube's muon track reconstruction**, 2011. 57, 58

[131] J. Brunner. **Antares simulation tools**. In *Antares simulation tools*. Presented at the VLVnT Workshop on Technical Aspects of a Very Large Volume Neutrino Telescope in the Mediterrenean Sea 2003, Amsterdam, The Netherlands, 2003. 58, 59

[132] Luis Alvarez and Luis Mazorra. **Signal and Image Restoration Using Shock Filters and Anisotropic Diffusion**. *SIAM J. Numer. Anal.*, **31**(2):590–605, April 1994. 59

[133] Pascal Getreuer. **Fast Gaussian Convolution with Recursive Filtering**. http://www.getreuer.info/home/gaussianiir, 2011. 59

[134] Valéry Brinnel. **Maximum likelihood optimization to improve IceCube's muon track reconstruction**. Bachelor thesis, Johannes Gutenberg-Universität Mainz, Mainz, Germany, 2012. 61

[135] M.G. Aartsen et al. **Energy Reconstruction Methods in the IceCube Neutrino Telescope**. *JINST*, **9**:P03009, 2014. 66, 67, 68

[136] Sandy Miarecki. private communication, 2012. 69

[137] T. Neunhöffer. **Die Entwicklung eines neuen Verfahrens zur Suche nach kosmischen Neutrino-Punktquellen mit dem AMANDA-Neutrinoteleskop**. *Berichte aus der Physik. Shaker*, 2003. 69, 70

[138] Steve Barnet. **IceCube: Evolving Work Flows**. In *IceCube: Evolving Work Flows*. Presented at the Paradyn/HTCondor Week 2013, Madison, Wisconsin, 2013. 73

[139] The IceCube Collaboration. **SLC hit cleaning**, May 2014. 73, 74

[140] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. Krasznahorkay, Jr., M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla. **TMVA - Toolkit for Multivariate Data Analysis**. *ArXiv Physics e-prints*, March 2007. 76, 77, 78, 80, 81

[141] Sirin Odrowski. private communication, 2013. 85

[142] Armen M. Atoyan and Charles D. Dermer. **Neutral beams from blazar jets**. *Astrophys.J.*, **586**:79–96, 2003. 93

[143] Charles D. Dermer, Kohta Murase, and Hajime Takami. **Variable Gamma-ray Emission Induced by Ultra-High Energy Neutral Beams: Application to 4C +21.35**. *Astrophys.J.*, **755**:147, 2012. 93, 94

[144] Joel R. Primack, Alberto Dominguez, Rudy C. Gilmore, and Rachel S. Somerville. **Extragalactic Background Light and Gamma-Ray Attenuation**. *AIP Conf.Proc.*, **1381**:72–83, 2011. 94

[145] A. Achterberg et al. **On the selection of AGN neutrino source candidates for a source stacking analysis with neutrino telescopes**. *Astropart.Phys.*, **26**:282–300, 2006. 94

[146] A. Yu. Neronov and D.V. Semikoz. **Which blazars are neutrino loud?** *Phys.Rev.*, **D66**:123003, 2002. 94

[147] Charles D. Dermer and Armen Atoyan. **Ultrahigh energy cosmic rays, cascade gamma-rays, and high-energy neutrinos from gamma-ray bursts**. *New J.Phys.*, **8**:122, 2006. 94

[148] Charles Dermer. private communication, 2012. 94

[149] Jim Braun, Jon Dumm, Francesco De Palma, Chad Finley, Albrecht Karle, et al. **Methods for point source analysis in high energy neutrino telescopes**. *Astropart.Phys.*, **29**:299–305, 2008. 96

[150] S. S. Wilks. **The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses**. *The Annals of Mathematical Statistics*, **9**(1):60–62, 03 1938. 97

[151] G. Ghisellini, L. Maraschi, and F. Tavecchio. **The Fermi blazars' divide**. *MNRAS*, **396**:L105–L109, June 2009. 100

[152] L. Maraschi, E. Pian, C. M. Urry, et al. **Multiwavelength Observations of a Dramatic High-Energy Flare in the Blazar 3C 279**. *Astrophysical Journal*, **497**:178–187, 1998. 106

[153] Inc ERI Economic Research Institute. **Table 7: Kolmogorov-Smirnov Test**. http://dlc.erieri.com/onlinetextbook/index.cfm ?fuseaction=textbook.appendix&FileName=Table7, 2013. 116