Unfolding the Atmospheric Neutrino Flux Using the IceCube/DeepCore Neutrino Telescope

Dissertation submitted for the award of the title

"Doctor of Natural Sciences"

to the Faculty of Physics, Mathematics, and Computer Science

of Johannes Gutenberg Universität Mainz

in Mainz, Germany

submitted by

Joakim Sandroos

Born on the 8th of May 1980 in Odense, Denmark
Mainz, November 29th, 2020

1. Gutachter:	Prof. Dr. Sebastian Böser
2. Gutachter:	Prof. Dr. <redacted>

Examined on April 1st, 2021

Joakim Sandroos
Experimentelle Teilchen- und Astroteilchen Physik
Institut für Physik
Staudingerweg 7
Johannes Gutenberg-Universität Mainz
D-55128 Mainz, Germany
jsandroo@uni-mainz.de

Abstract

Cosmic rays in the form of high energy particles, strike the Earth every second. Since the first discoveries of the cosmic ray flux in the early nineteenth century, the origin and properties of the cosmic rays have been the subject of much investigation and debate. Cosmic ray interactions in the upper atmosphere create particle showers, containing several species of particles. The daughter particles generated in these showers carry with them information about the interaction dynamics and properties of the parent cosmic rays. One group of daughter particles are the neutrinos, which are left handed leptons included in the standard model of particle physics. While they were previously believed to be massless, progress over the past two decades have established their mass through observation of oscillations, particularly in atmospheric neutrinos. This work presents an investigation of the atmospheric neutrino flux measured with the IceCube South Pole Neutrino Observatory, using data from 5 seasons of IceCube operation. Modeling of the atmospheric neutrino flux is a multi dimensional problem and at the time of publication the uncertainties on both theoretical predictions and measurements are of such scale that all models and measurements are in agreement. This work aims to improve the precision in measurement in order to give better discriminating power between models. The analysis presented herein performs an unfolding in three dimensions, energy, zenith angle and particle identification channel. The method is as model independent as possible and utilizes iterative Bayesian unfolding to measure the unfolded event rate by detection volume for two particle groups: $\nu_\mu^{\mathrm{cc}} + \bar{\nu}_\mu^{\mathrm{cc}}$ constituting the muon neutrino charged current interactions, and ν_{rest} constituting all other flavor and interaction types. A total of 204847 neutrino candidate events are observed and unfolded. Great care is taken to avoid bias in the unfolding via a series of closure tests. The unfolded results yield true interaction rates on the scale of $10^{-10}[\mathrm{m}^{-3}\mathrm{s}^{-1}]$. The unfolded results are generally in good agreement with previous unfolded data from Super Kamiokande, but do show some tension with current predictions - particularly in the low energy region below 80GeV and in the upgoing direction. The relative uncertainty on the unfolded result lies between 3% and 13% for the energy spectrum, with low uncertainty particularly in the region of interest to neutrino oscillation measurements. For the $\cos(\theta_z)$ spectrum the relative uncertainty lies between 4% and 7.5% in the upgoing region, while it rises to 18.5% in the downgoing region. Even with these uncertainties, the results presented in this dissertation constitute the most precise measurement of the atmospheric neutrino flux at the time of publication.

Contents

List of Figures

List of Tables

1 Introduction and Motivation

Atmospheric neutrinos are elementary particles generated by cosmic ray interactions in the atmosphere. The main analysis goal of this dissertation is an atmospheric neutrino measurement, using the *IceCube South Pole Neutrino Observatory*. This will be performed in as model independent a way as possible. Any such measurement requires an amount of background knowledge: both of atmospheric neutrinos and cosmic rays; but also of neutrino oscillations and neutrino interactions. These topics will be discussed both in the coming sections of this chapter and in chapter 2. The remaining chapters will cover the IceCube detector in chapter 3, the data sample and simulation used in chapter 4, the Bayesian iterative unfolding analysis method in chapter 5, statistical and systematic uncertainties in chapters 6 and 7 before presenting the analysis and results in chapter 8. The final chapter, 9, discusses the results, relates them to other measurements in the field and gives an outlook on continuation of this work.

1.1 Consistency Between Measurement and Theory

Models of the neutrino flux have associated uncertainties on the scale of about 10% [7], while modern flux measurements typically come with errors on the scale of about 20% [3]. Within these errors, the experimental data and the flux predictions agree. However, the difference between various flux models is much smaller than the scale of the uncertainties making distinction between models impossible. It is thus important, in order to bring the field forward, to reduce both the uncertainty in the models and the measurement error of the data. In this work the aim is to contribute the most precise measurement of the atmospheric neutrinos to date, and release the data in way which is meaningful and useful for the neutrino community at large.

1.2 Current Knowledge of the Atmospheric Neutrino Flux

The atmospheric neutrino flux has been measured by various experiments at several different locations on Earth over the last three decades. Despite this, the knowledge of the flux is still associated with large uncertainties up to about 100% relative uncertainty and a low level of resolution in the energy spectrum. The major contributions to the knowledge of the atmospheric neutrino flux are shown in fig. 1.1 [3]. The earliest measurements listed date from the Frejus experiment in the 1990s to the most recent measurement by Super Kamiokande in 2016 [3]. As can be seen in fig. 1.1, the muon neutrino spectrum has been investigated by several experiments in both the high energy region above 1TeV, as well as in the low energy region below $\sim$ 10 GeV. Both of these regions show several data points. However, the energy band between 10GeV and 100GeV shows some sparsity in the measurement data, which is unfortunate as this is also the main region of interest to neutrino oscillation experiments. Another goal of this work is to provide a measurement with a dense binning in exactly this region of the energy spectrum.

1.3 Goals of this Work

For clarity the goals of this work will be summarized here. Based on the discussion of the current status of atmospheric neutrino measurements, the work presented in this thesis has several goals: Firstly, to improve the measurement precision over previous efforts, and thus bring the uncertainty down. Secondly, to provide measurement data in the sparsely populated region of interest to oscillation experiments between 10GeV and 100GeV. Thirdly, to perform the measurement in an as model

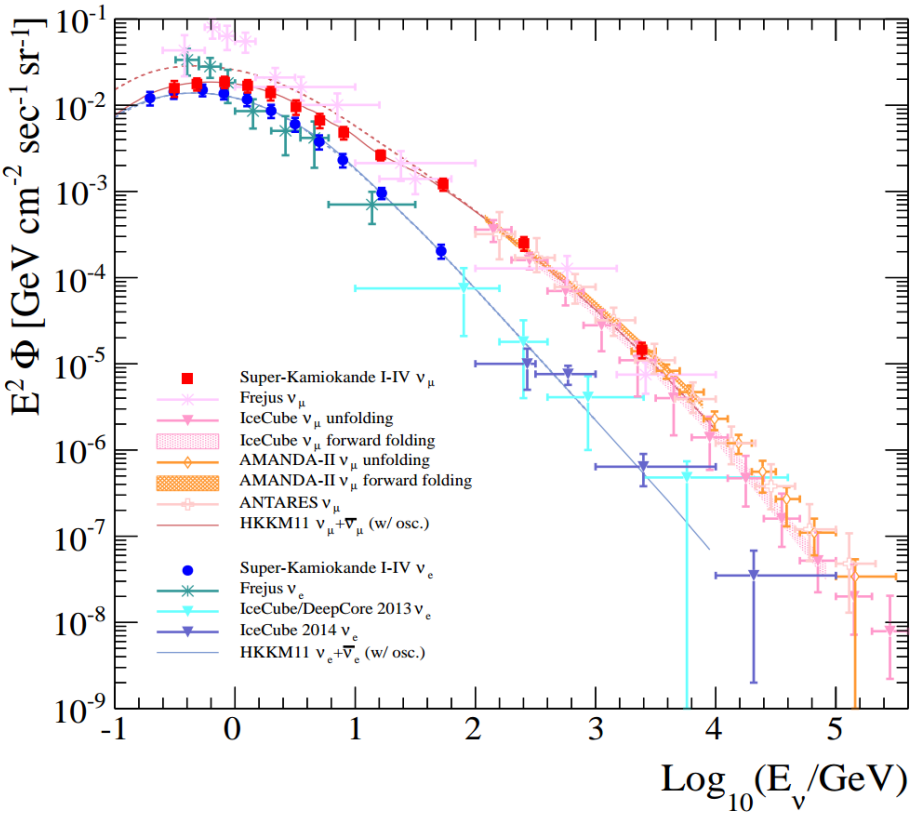


Figure 1.1: Historic overview of global neutrino flux measurements, for the ν_μ and ν_e channels. Neutrino flux as a function of energy. The flux is weighted by E^2 to better illustrate the shape of the energy spectrum. Both the high- and low energy regions are highly populated, but the region from 10GeV to 100GeV is sparse in data points. Figure from [3].

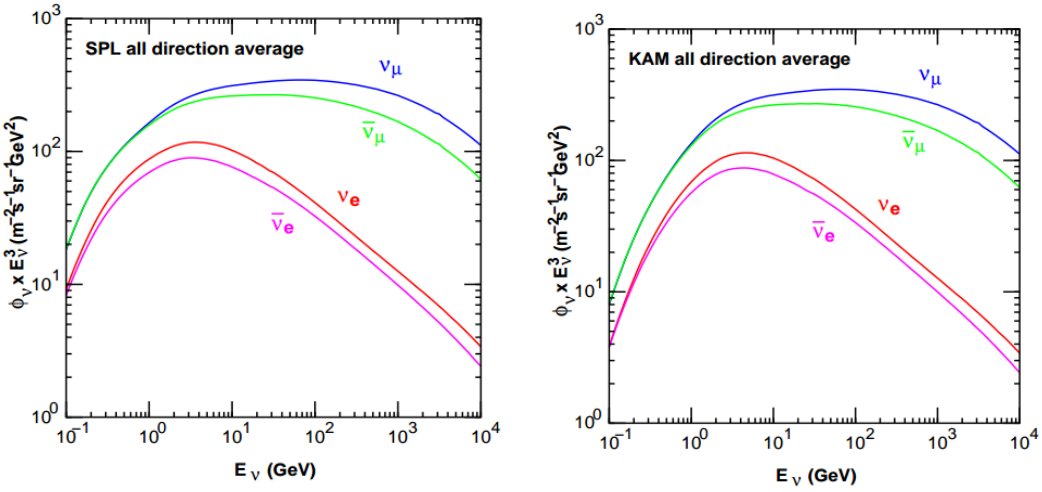


Figure 1.2: HKKM Atmospheric neutrino flux energy spectrum prediction as simulated in the HKKM model, weighted by E^3 so that spectrum features are visible. Locations are the South Pole (left) and Kamioka (right). Graph from [7].

independent way as possible, so a data release can be easily used by model builders without prior knowledge of the IceCube experiment.

1.4 Modeling the Atmospheric Neutrino Flux

In recent decades much effort has gone into modeling the atmospheric neutrino flux. Two major simulations exist in the field: The *Matrix Cascade Equation* solver (MCEq) by A. Fedynitch [8], and the *Honda-Kajita-Kasahara-Midorikawa* model (HKKM) by M. Honda et al. [7] (colloquially known as the 'Honda' model). These models rely on measurements or assumptions relating to the cosmic ray (CR) primary flux, the interaction cross sections of CR on molecules in the air, and the density of the atmosphere. The primary CR interactions give rise to a *cascade* of daughter particles in the atmosphere, sometimes referred to as an *air shower*. This leads to both a neutrino flux and to a muon counterpart. The properties of these air showers such as energy, directionality and whether they are mostly hadronic or electromagnetic, leave a characteristic imprint on the neutrino flux. As such, many factors need to be taken into account when simulating the atmospheric neutrinos: densities, cross sections, magnetic field strength, and flavour ratios are amongst the most important parameters. The atmospheric neutrino flux is expected display a seasonal variation induced by the temperature dependent atmospheric density.

MCEq works by solving the cascade equation discussed in section 2.4, in one dimension, meaning it is accurate only insofar as daughter particles are emitted predominantly in the forward direction. HKKM on the other hand is fully three dimensional and tracks simulated CR from generation, through propagation in the Earth's magnetic field, to interaction in the atmosphere. The initial CR flux is calibrated by measurements from the AMS02 experiment on board the international space station [9]. The atmospheric neutrino flux at the South Pole, as predicted by the HKKM model, is shown in figures 1.2 and 1.3. Due to the geometry of the Earth and because the scattering not being as forward

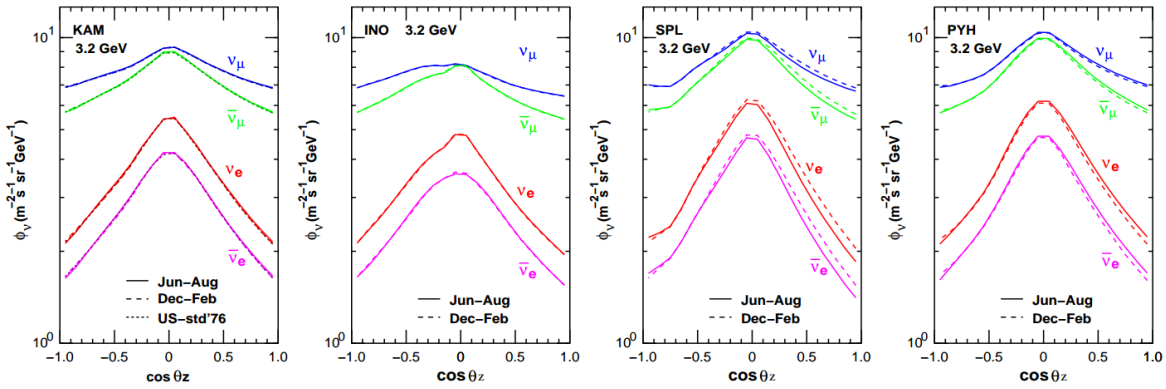


Figure 1.3: HKKM model predictions for the atmospheric neutrino flux arrival direction in zenith angle at 3.2 GeV energy. The locations are, left to right: Kamioka, Indian Neutrino Observatory, South Pole, and Pyhasalmi. Notice the increase around the horizon, which is caused by a geometric effect. The solid and dashed lines are for the summer months and winter months respectively, which illustrates the seasonal variation of the flux. Graph from [7].

directed in the energy regime below $\sim$ 5GeV as for higher energies, the atmospheric neutrino flux is expected to show a peak at zenith angles around the horizon [10]. Using the HKKM model simulation files, a direct comparison of the muon neutrino flux between Kamioka and the South Pole is shown in fig. 1.4.

1.5 Why Unfold?

In many cases in particle physics, performing an unfolding of the measured data is not necessary - particularly in the cases where an analysis seeks to compare a model prediction to data via one or more parameters. One could simply modify the model prediction to include distortions due to detector effects and proceed to compare prediction directly against measurement via statistical tests. This is the traditional way of performing a forward folding analysis, fitting for physical parameters of interest. However, such a measurement leaves no room to compare experimental data directly between different experiments, because the detector responses vary between experiments. Comparison can then only be done on the outcome of the statistical tests, usually represented as confidence intervals on physical model parameters. This method requires a paradigm in which model testers agree on which models and parameters to test for, in order for the comparisons to be meaningful. Unfolding overcomes this challenge by answering a fundamentally different question from the forward fitting models. Whereas traditional forward fitting yields results based on assumptions of a model and the parameters contained therein, unfolding aims to yield results based on knowledge of the experimental setup. Where forward fitting provides the most likely values for the parameters of a model, unfolding provides the most likely data set to yield the actual measurement. Thus, results from the two methods can not be compared directly, and they should not be thought of as competing: each method has its raison d'etre and justification. In this work the aim is to provide a 'true' data sample that can be used as input for other projects, which is the topic of the next section.

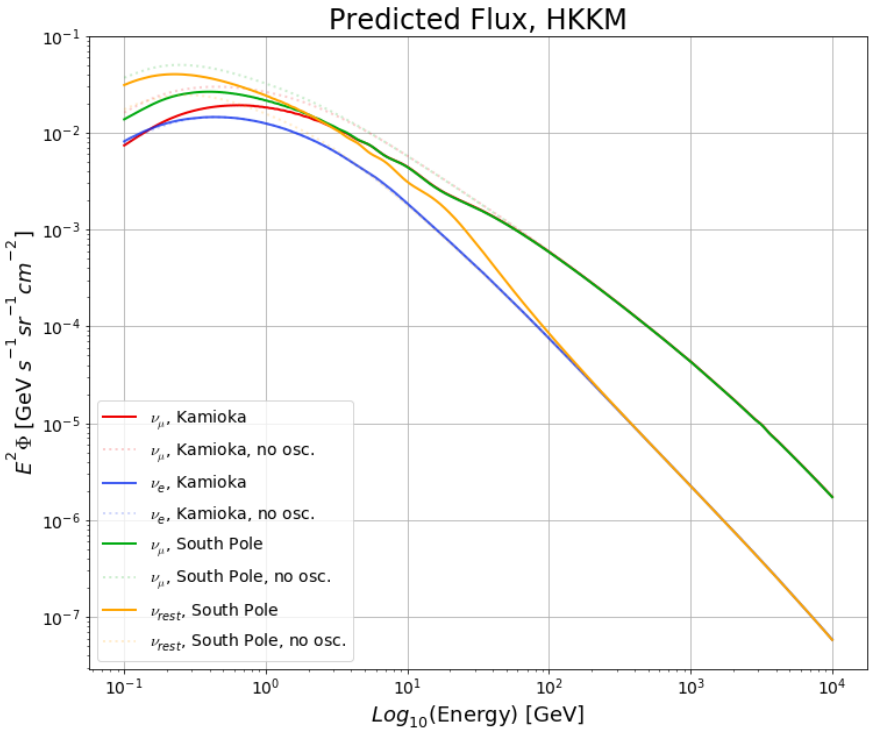


Figure 1.4: Flux prediction from the HKKM model for Kamioka and the South Pole. The solid and dashed lines show the oscillated and un-oscillated fluxes, respectively. The cross over between the ν_μ and ν_{rest} curves are caused by inclusion of the ν_τ component. The deviation at energies below $\sim$ 30GeV stems from the latitude effect described in chap. 2.3

1.6 Unfolded Data as Input to Other Measurements

The atmospheric neutrino flux is taken as input to many neutrino experiments: Super Kamiokande, SNO, Minos, ICARUS, Antares, Frejus, KM3NeT, IceCube, and several others. The measurements performed by all of these experiments are limited in precision by the current best knowledge of the flux. As such a solid understanding of the atmospheric neutrino flux is essential, not simply from a basic science point of view, but as a stepping stone toward more precise measurements of other parameters in areas such as neutrino oscillations or mass hierarchy measurements. However, the measured quantity in the experiments depend not only on the flux, but also on the interaction cross section of the neutrinos in the detection medium, that is also associated with considerable uncertainty. Therefore, measurements of the flux must make assumptions regarding the cross sections and vice versa, as both are associated with uncertainties. The interesting quantity for an experiment aiming to measure an oscillation signal thus becomes a convolution of the flux and the cross section, which can be described as an event rate. A model independent measurement of the true neutrino rate, a convolution of the flux and the cross section, in IceCube is the main goal of the analysis presented in this thesis. Such a data set would be of great interest to any neutrino experiment relying on atmospheric neutrinos.

2 Neutrino and Cosmic Ray Physics

This chapter focuses on a discussion of neutrino physics from a theoretical as well as historical perspective. The following sections will give a historical overview of the theoretical background and experimental progress leading to the modern understanding of neutrino physics.

2.1 The Standard Model

The standard model of particle physics, developed mainly during the nineteen sixties and seventies, describes the elementary particles in our universe as well as their interactions. The Standard Model is well described in scientific literature [12, 13] and the *Particle Data Group*, an international collaboration [14], publishes yearly reviews of the experimental evidence supporting the theory in the "*Review of Particle Physics*" [6]. It describes 17 particle types and groups them into *fermion* and *boson* categories as shown in fig. 2.1. The fermions consist of quarks and leptons and are subdivided into three generations. These make up all known matter in the universe. The bosons are divided into vector bosons and scalar bosons. The vector bosons are force carriers and correspond to the basic forces of nature: Electromagnetism, mediated by the photon, the weak nuclear force, mediated by the W and Z bosons, and the strong nuclear force mediated by the gluons. The force of gravity is not included in the standard model per se, although a mass generating mechanism is included via a scalar boson, known as the Higgs boson. The model has been tremendously successful in describing the quantum world of particles. It predicted the existence of several particles, which were subsequently discovered in experiments. All of these were significant discoveries in their own right, but one deserves a special mention: The discovery of the Higgs boson in 2012, solved a large part of the mystery surrounding the origin of particle mass, and with it all particles predicted by the standard model of particle physics had been discovered. A schematic overview of the known particles is shown in fig. 2.1. Even with this tremendous success of the standard model, the description of nature is not complete and there are several outstanding challenges such as the baryon asymmetry, dark matter and neutrino mass. In the standard model, neutrinos are modeled as massless fermions of left handed chirality. However over the past several decades neutrino mass has been firmly established experimentally via measurements of neutrino oscillations. While it is possible to extend the standard model in order to include a standard mass term via a coupling to the Higgs boson, this would then require the existence of a right handed neutrino type. As these would be right handed they would not participate in the weak interaction, making direct detection difficult. Much effort in recent years have gone into investigating the possibility of the existence of right handed neutrinos. Particularly atmospherically generated neutrinos have been utilized for oscillation measurements, giving another reason for the atmospheric neutrino spectrum to be the target of measurement in its own right. The atmospheric neutrinos stem from cosmic ray interactions in the Earth's atmosphere, which is the topic of the next section.

2.2 Cosmic Rays

Cosmic rays have been known since 1912, when Victor Hess did his famous balloon borne experiments [4, 5]. Unexplained radiation had been detected and it was hypothesized that it originated in the Earth. To much surprise, Hess' measurements showed the exact opposite: when he moved his detectors higher in the atmosphere, the event rates increased, indicating the radiation originated in or beyond the atmosphere. It was dubbed *cosmic rays*. Many experiments have since investigated the properties of cosmic rays, and an illustrative figure of the flux scaled by $E^{2.5}$ is shown in fig. 2.2. A more recent version from the International Cosmic Ray Conference 2019, summarizing many recent results, is found

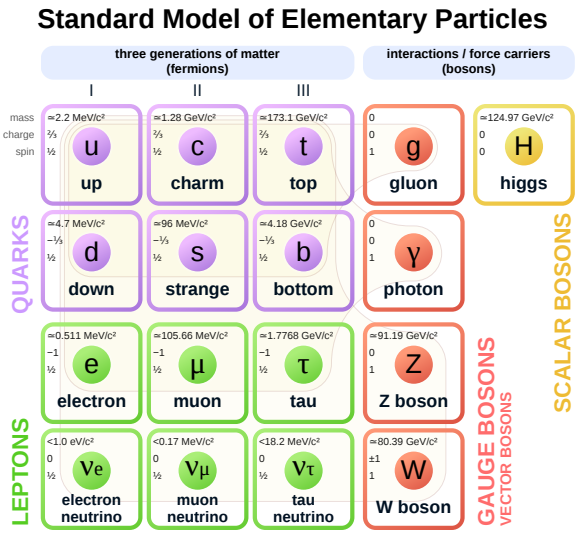


Figure 2.1: Schematic of particles contained in the standard model of particle physics. Figure from [15].

in fig. 2.3 [19]. The basic structure is the same between them: a steeply falling spectrum, showing a few distinct features. Most notably, the break in the spectrum at about 3×10^{15}eV/particle in fig. 2.2, known as 'the knee', could indicate a broken power law dependence. With more data the existence of this break has been confirmed and as measurements proceeded to higher energies, a second break was found, dubbed 'the ankle'. These breaks in the spectrum indicate contributions from different sources of acceleration, and a full understanding of the source composition is an ongoing effort in the field. The cosmic ray composition has also been investigated and found to be made up mostly of Hydrogen nuclei, as illustrated in fig. 2.4. Care must also be taken to account for the charged particles in the Cosmic Rays bending in the Earth's magnetic field, which is covered by the next section.

2.3 Geomagnetic Effects

Cosmic rays carry positive magnetic charge as they are made up of atomic nuclei, predominantly Hydrogen. The cosmic ray composition is illustrated in fig. **??** [6]. Having electric charge, the cosmic ray's trajectory will bend in Earth's magnetic field. Consider a particle of charge $Z \cdot e$ propagating with velocity $\vec{v}$ in the equatorial plane of a magnetic dipole field $\vec{B}$, as illustrated in fig. 2.5. The equilibrium condition is found when equating the Lorenz force $\vec{L}$ with the centrifugal force $\vec{c}$:

$$Ze|\vec{v} \times \vec{B}| = \frac{mv^2}{r} \tag{2.1}$$

The consequence of the cross product behaviour is interesting and has physical consequences for the flux: As the magnetic field strength increases, the curvature radius will decrease, under assumption of constant velocity. In essence: The closer a charged particle gets to the surface, the stronger repellent effect of the magnetic field. The critical momentum at which a particle can reach the Earth, The

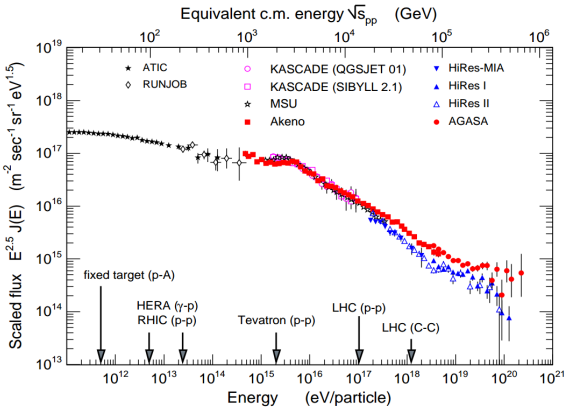


Figure 2.2: Cosmic ray energy spectrum spanning several orders of magnitude in both energy and flux. The energy per nucleon of the incoming cosmic ray is shown on the x-axis, with the energy of several collider experiments also marked for comparison. The upper x-axis shows the equivalent center of mass energy for the cosmic ray interactions. The difference between the two axes stems from CR interactions corresponding to fixed target experiments. Graph from [16].

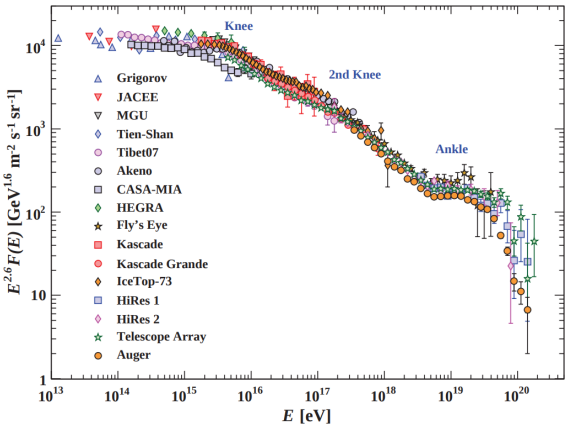


Figure 2.3: Recent summary of cosmic ray energy spectrum measurements from a range of experiments. The presence of the breaks known as the 'Knee' and the 'Ankle' are clearly visible. Furthermore, indications of a '2nd Knee' can be seen at around 10^{17}eV Figure from [6].

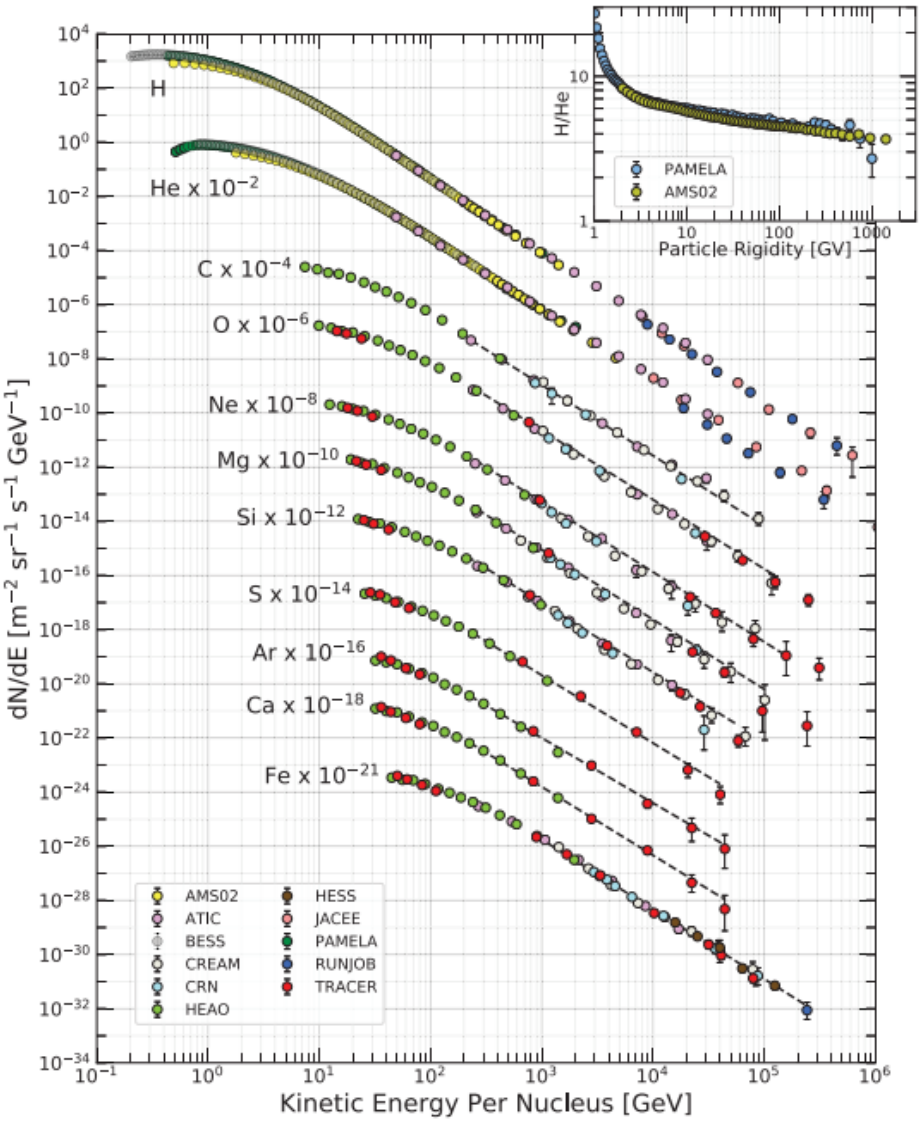


Figure 2.4: Fluxes of nuclei of the cosmic ray primary particles in number of particles per energy per nucleus as a function of energy per nucleus. The inset at the top right shows the H/He ratio at constant rigidity. Figure from [6].

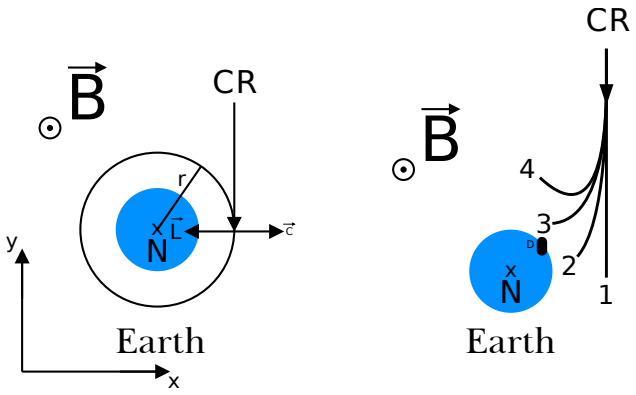


Figure 2.5: Illustrating the geomagnetic effect on incident cosmic rays. The Earth is observed along the magnetic North/South axis, from the North, meaning the magnetic field lines at the equator point out of the page. Left: Vector diagram of forces. r is the curvature radius along which the Lorentz force and the centrifugal force cancel out. Right: CRs with progressively lower energies follow trajectories marked 1 through 4. As the energy decreases, the bending of the trajectory in the magnetic field increases, giving rise to the momentum cut-off. The magnetic field is responsible for the east/west asymmetry due to the broken symmetry of the Lorentz force.

momentum cut-off per proton $\frac{p}{Z}$, can be calculated by using $B = \frac{\mu_0 M}{4\pi r^3}$, setting $r = R_\oplus$ and using the magnetic moment $M = 8 \times 10^{22}$Am:

$$\frac{p}{Z} = \frac{\mu_0}{4\pi} \frac{eM}{R_\oplus^2} \approx 59.6 GeV \tag{2.2}$$

A few comments are in order: Firstly, the effect decreases as the angle between $\vec{v}$ and $\vec{B}$ approaches zero, predicting the CR flux to be larger at the poles than at the equator. Secondly, the cross product dictates the direction of the Lorenz force based on the directionality of the Earth's magnetic field - leading particles to always be bent stronger in the eastern direction. This predicts a stronger CR flux from the west as compared to the east and gives rise to an east/west asymmetry. Both the latitude effect and the east/west asymmetry have been verified experimentally as shown in fig. 2.6 and 2.7 [17, 18]. A separate contribution to the latitude effect is the density of the atmosphere which is temperature dependent, and thus also depends on the latitude, at which measurements are done. This effect is illustrated in the upper panel of fig. 2.6. The characteristic features of the cosmic ray spectrum, will be inherited to the neutrinos generated by the cosmic ray interactions in the Earth's atmosphere. The development of the atmospheric shower is described by the *cascade equation* which is the subject of the next section.

2.4 The Cascade Equation

The Earth is constantly bombarded with radiation from the universe. These Cosmic rays incident on the Earth interact in the atmosphere, generating pions and kaons, leading to neutrinos as they in turn decay. The primary interaction usually takes place at a typical altitude of 15-20 km [7], depending

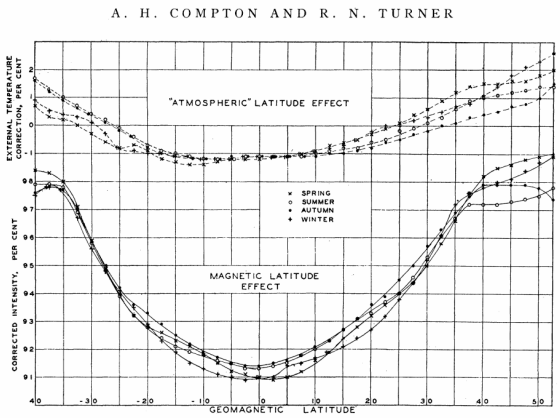


Figure 2.6: Magnetic (solid) and atmospheric (dashed) latitude effect on the cosmic ray flux. Figure from [17]

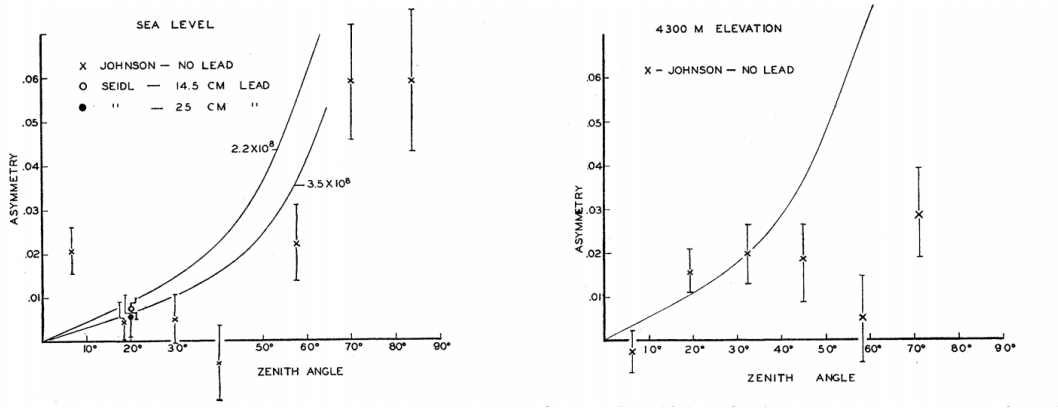


Figure 2.7: East/west asymmetry as a function of zenith angle. Latitude at sea level (left) and at 3400m above sea level (right). The curve shows the theoretical prediction based on an E^{-3} power spectrum. Figure from [18]

on factors such as atmospheric density and energy of the incoming cosmic ray. From the interaction vertex a cascade of daughter particles propagate downward, and the flux of particles in the atmosphere can be described by the cascade equation [35]:

$$\frac{dN_i(E_i,X)}{dX} = -\frac{N_i(E_i,X)}{\lambda_i} - \frac{N_i(E_i,X)}{d_i} + \sum_{j=i}^{J} \int_E^\infty \frac{F_{ji}(E_i,E_j)}{E_i} \frac{N_j(E_j,X)}{\lambda_j} dE_j \tag{2.3}$$

Here $N_i(E_i,X)$ represents the flux of particles of type i with energy in the range E to $E+dE$. X marks the slant depth in the atmosphere, which is the length along the trajectory traversed by the particle in atmosphere. The constants λ_i and d_i represent the average interaction length and the average decay length respectively. As such the first two RHS terms govern the loss of particles due to these processes. The third term represents production of particles of type i, and the function $F_{ji}(E_i,E_j)$ is the transfer function describing the dimensionless particle yield following from an inclusive cross section for a particle of type j to interact with a nucleus in the atmosphere and produce an outgoing particle type i with energy: $E_i < E_j$ [35]:

$$F_{ji}(E_i,E_j) \equiv E_i \frac{1}{\sigma_j^{\text{air}}} \frac{d\sigma_{j\text{air}\to i}}{dE_i} = E_i \frac{dn_i(E_i,E_j)}{dE_i} \tag{2.4}$$

dn_i is then the average number of particles of type i produced in bin $E+dE$ for each interaction of particle type j. Gaiser et al in [35] gives an important example on page 109 where the authors define the spectrum-weighted moment:

$$Z_{ab} \equiv \int_0^1 \chi_L^\gamma \frac{dn_b}{d\chi_L} d\chi_L, \tag{2.5}$$

where χ_L is the fraction: $E_{\text{lab}}/E_{\text{beam}}$. The spectrum-weighted moment is used to calculate the pion spectrum arising from a power law spectrum of nucleons (N) passing through a target of thickness dX. The nucleon power law spectrum is defined as:

$$N(E) = KE^{-(\gamma+1)}, \tag{2.6}$$

and the pion spectrum can then be calculated as:

$$\frac{d\Pi(E_\pi)}{dX} = \frac{N(E_\pi)}{\lambda_N} Z_{N\pi} \tag{2.7}$$

A few comments are in order. Firstly: when formulated like this, the spectrum-weighted moment Z_{ab} has the advantage of directly indicating the strength of the particle production given the spectrum $N(E_\pi)$. Secondly: there is no change in the index of the power law, meaning the spectrum of the daughter particles (in this example, pions) will have the same spectral index as the beam spectrum. This makes the power law index a useful 'fingerprint' when investigating distributions and interactions in astrophysics. Using the cascade equation it is possible to estimate the production of kaons and pions from the incoming cosmic rays.

2.5 Atmospheric Neutrinos from Cosmic Rays

The Kaons and Pions deposit their energy in the atmosphere. Neutrinos are generated as an incoming cosmic ray with atomic number A_{cr} strikes an atom in the atmosphere with atomic number A_{atm}. The main neutrino generating processes can be written as follows [35]:

Particle	Result	Mode	Branching Ratio
K^+	$\mu^+ + \nu_\mu$	leptonic	$63.56\% \pm 0.11\%$
K^+	$\pi^+ + \pi^0$	hadronic	$20.67\% \pm 0.08\%$
K^+	$\pi^+\pi^+\pi^-$	hadronic	$5.583\% \pm 0.024\%$
K^+	$\pi^+\pi^0\pi^0$	hadronic	$1.760\% \pm 0.023\%$
K^+	$\pi^0 + e + \nu_e$	semileptonic	$5.07\% \pm 0.04\%$
K^+	$\pi^0 + \mu^+ + \nu_\mu$	semileptonic	$3.352\% \pm 0.033\%$
K^0_S	$\pi^+\pi^-$	hadronic	$69.20\% \pm 0.05\%$
K^0_L	$\pi^\mp e^\pm \nu_e(\bar{\nu}_e)$	semileptonic	$40.55\% \pm 0.11\%$
K^0_L	$\pi^\mp \mu^\pm \nu_\mu(\bar{\nu}_\mu)$	semileptonic	$27.04\% \pm 0.07\%$

Table 2.1: Prominent Kaon decay channels and their branching ratios (>% level).

$$A_{\rm cr} + A_{\rm atm} \rightarrow \pi^\pm + ... \tag{2.8}$$
$$\hookrightarrow \mu^\pm + \nu_\mu(\bar{\nu}_\mu) \tag{2.9}$$
$$\hookrightarrow e^\pm + \nu_e(\bar{\nu}_e) + \bar{\nu}_\mu(\nu_\mu) \tag{2.10}$$

A similar chain exists for $A_{\rm cr} + A_{\rm atm} \rightarrow K^\pm$. While the branching ratio for $\pi^\pm \rightarrow \mu^\pm + \nu_\mu(\bar{\nu}_\mu)$ is 100%, the kaon decay only results in a direct daughter muon in approximately 63.5% of the cases [35]. However, several other kaon decay modes can contribute to the neutrino content of an atmospheric cascade. The kaon decay channels and their branching ratios are listed in table 2.1, for channels with branching ratios above the 1% level. Most of the branches contain pions, which have an average lifetime of about $2.8 \times 10^{-8} s$, but due to length contraction as seen from the rest frame can propagate long distances in the atmosphere. Due to this effect a large fraction of secondary pion decays at high energies can be safely ignored. While $(K^0_S + K^0_L)$ have high branching ratios into final states containing neutrinos, these kaons contribute little to the overall neutrino flux as they are produced in a similar number to K^+.

In the cases where all daughter particles decay it is thus reasonable to expect the ratio of muon neutrinos to electron neutrinos to be around two:

$$\frac{(\nu_\mu + \bar{\nu}_\mu)}{(\nu_e + \bar{\nu}_e)} \sim 2 \tag{2.11}$$

and the ratio of neutrinos to anti neutrinos to be similar to the ratio of anti muons and muons:

$$\frac{\nu_e}{\bar{\nu}_e} \sim \frac{\mu^+}{\mu^-} \tag{2.12}$$

A few comments are appropriate at this point. Due to time dilation the mean decay length of the muons can become longer than the average production height. This means a significant fraction of muons would simply pass by detector locations before decaying, lowering the incident number of electron neutrinos. Any detector should expect the ratios of $\nu_e/\nu_\mu$to drop at energies above approximately one GeV. The first detections of atmospheric neutrinos from cosmic ray interactions came from two independent experiments in 1965: the 'Kolar' experiment in India and the 'South African neutrino experiment', located at 3.2 km depth in a mine in South Africa [36]. The latter reported the

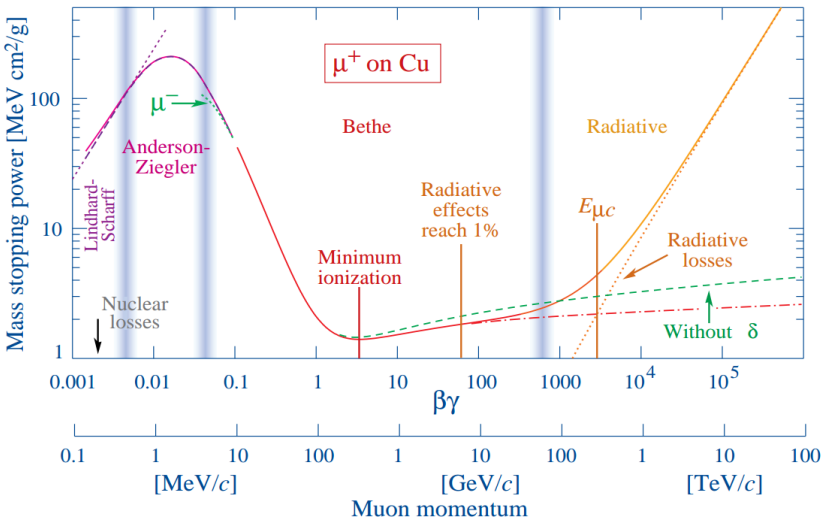


Figure 2.8: Illustrating the mass stopping power of positive muons as a function of momentum. Figure from [6].

detection muons produced from neutrino interactions yielding 22 events over a full detector equivalent live time of $\sim$ 94 days vs an expectation of less than one event per year from atmospheric muons. Later two other experiments were able to confirm the directionality of the muons to indeed be in the upgoing direction, which would simply be impossible for an ionizing muon, due to energy loss during propagation, which is the topic of the next section.

2.6 Charged Particle Energy Loss

As described in the previous section, CR interactions in the atmosphere generate a host of daughter particles. An important component is the muon, which is capable of propagating distances spanning several kilometers and can induce Cherenkov radiation. Atmospheric muons constitute an important background for any measurement utilizing atmospheric neutrinos. This background is discussed in more detail in sec. 4.4, here the muon properties most relevant for IceCube are discussed.

Energy loss during propagation occurs for all charged particles via electronic interactions. These are single collisions with energy loss and lead to ionization or atomic excitations of the interaction material [6]. Typically the energy loss is small compared to the energy of the particle - below 100eV in 90% of cases. For free electrons the interactions are adequately described by Rutherford scattering, however, in matter the electrons are not free. The energy transfer between the charged particle and the electron, W, must be finite and depends on the atomic structure of the material. Bethe obtained the differential cross section by introducing a correction $B(W)$ to the Rutherford differential cross section $\frac{d\sigma_R(W;\beta)}{dW}$ [6]:

$$\frac{d\sigma_B(W;\beta)}{dW} = \frac{d\sigma_R(W;\beta)}{dW} B(W) \tag{2.13}$$

This equation, however, is only valid in the energy regime in which atomic interactions are important - above such energies, it does not provide a reasonable description of nature, and the standard Rutherford differential cross section can be used. As charged particles scatter in the medium, they transfer energy to the medium and thus the mean energy loss is given by the Bethe equation [6]:

$$\left\langle -\frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right], \tag{2.14}$$

where Z and A are the proton and mass numbers of the atoms making up the material and β and γ are defined as in special relativity. $\delta(\beta\gamma)$ is a correction due to the density of the material. The differential energy loss from eq. 2.14 is also known as the stopping power, and is illustrated for anti muons in copper in fig. 2.8. In the region of momentum between $\sim$ 1GeV to $\sim$ 100GeV the muons are considered to be minimally ionizing as radiative effects contribute less than 1% to the overall energy loss. In this work neutrinos with energies less than 1TeV are considered, so any secondary muons from neutrino interactions in the Antarctic ice can be safely assumed to be minimally ionizing. A muon or electron from such an interaction will lose about 220 MeV/m meaning it is reasonable to expect muonic Cherenkov track lengths to be predominantly in the range below 100m with a tail extending up to the km scale.

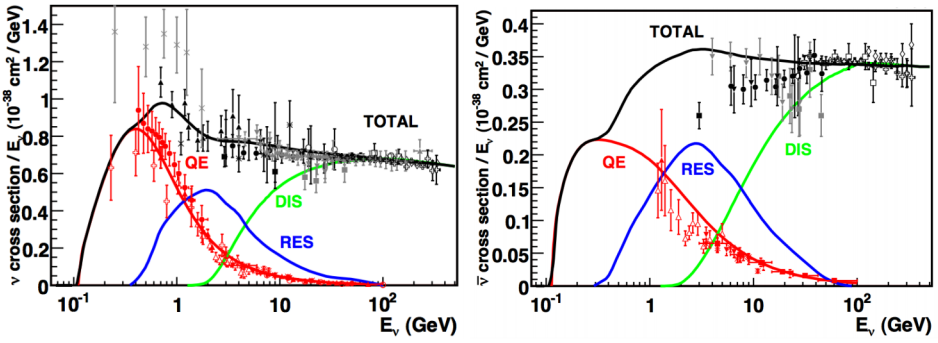


Figure 2.9: Muon neutrino (left) and antineutrino (right) charged-current cross section measurements and predictions. The contributing processes are in order of energy regime: quasi elastic (red), resonance production (blue) and deep inelastic scattering (red). Figure from [38].

Reaction	Equation	Boson
Electron emission	${}^{A}_{Z}X \rightarrow {}^{A}_{Z+1}X + e^- + \bar{\nu}_e$	W
Positron emission	${}^{A}_{Z}X \rightarrow {}^{A}_{Z-1}X + e^+ + \nu_e$	W
Electron capture	${}^{A}_{Z}X + e^- \rightarrow {}^{A}_{Z-1}X + \nu_e$	W
Positron capture	${}^{A}_{Z}X + e^+ \rightarrow {}^{A}_{Z+1}X + \bar{\nu}_e$	W
Electron annihilation	$e^- + e^+ \rightarrow \nu_e + \bar{\nu}_e$	W
Electron annihilation	$e^- + e^+ \rightarrow \nu + \bar{\nu}$	Z
Neutrino capture	${}^{A}_{Z}X + \overset{(-)}{\nu}_e \rightarrow {}^{A}_{Z\mp1}X + e^{\pm}$	W
$e^-\nu$ scattering	$e^- + \overset{(-)}{\nu}_e \rightarrow e^- + \overset{(-)}{\nu}_e$	W
$e^-\nu$ scattering	$e^{\pm} + \overset{(-)}{\nu}_e \rightarrow e^{\pm} + \overset{(-)}{\nu}_e$	Z
Neutrino scattering	${}^{A}_{Z}X + \overset{(-)}{\nu} \rightarrow {}^{A}_{Z}X + \overset{(-)}{\nu}$	Z

Table 2.2: Table of neutrino interactions and the mediating bosons. [6]

2.7 Neutrino Interactions

Neutrinos are known to be light left-handed spin-$\frac{1}{2}$ leptons that couple to quarks and leptons via the weak force mediated by the $W^{\pm}, Z^0$ bosons. Neutrinos are massive and so will also interact gravitationally, although gravity lacks a description in the standard model of particle physics. While neutrino interactions can occur in different scenarios, this section will focus on the scenarios and energies relevant to neutrino detection in ice: Neutrino interactions on nucleons and neutrino-electron scatterings. In terms of interactions the couplings are classified as *charged-current* interactions (cc) in which a W-boson is exchanged, or *neutral current* interaction (nc), mediated by the Z^0-boson. Being neutral in electric charge, neutrinos can only be detected via secondary particles stemming from their interactions in a medium. In broad terms these can be grouped into 5 distinct scenarios: *quasi-elastic scattering* (QE), *resonant scattering* (RES), *deep inelastic scattering* (DIS), *coherent scattering* (COH) and *neutrino-electron scattering* (NES). These sub-categories, based on the interaction kinematics, are illustrated in fig. 2.10. QE, RES, and DIS interaction each dominate in a specific energy regime, as illustrated in fig. 2.9 [37]. The following presents a brief overview of these three interactions.

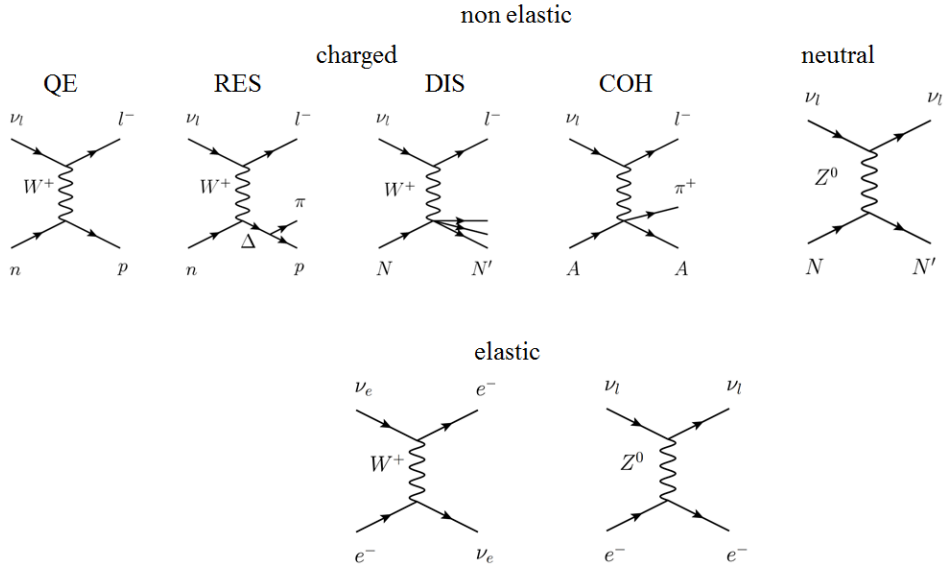


Figure 2.10: Feynmann diagrams of the most relevant neutrino interactions

Quasi-Elastic Scattering

In this scenario the incident neutrino scatters on an entire nucleon as opposed to one of its constituent partons. For the charged current interaction, these are the typical inverse beta decay reactions utilized for neutrino detection in many experiments:

$$\nu_e + n \rightarrow e^- + p^+, \qquad \bar{\nu}_e + p \rightarrow n + e^+ \tag{2.15}$$

The quasi-elastic differential cross section can be expressed as:

$$\frac{d\sigma}{dQ^2} = \frac{G_F^2 M^2 |V_{ud}|^2}{8\pi E_\nu^2} \left[A \pm \frac{(s-u)}{M^2} B + \frac{(s-u)^2}{M^4} C \right] \tag{2.16}$$

Resonant Scattering

The resonant scattering process produces single pions from inelastic scattering of neutrinos on nucleons. With enough energy the incident neutrino can excite the struck nucleon into an excited state. At energies between 1-4 GeV this is the dominant mode of neutrino-nucleon interaction as illustrated in fig. 2.9. This is due to the $\frac{1}{s-m^2}$ behaviour in the propagator, leading to a sharp increase in cross section at energies close to the mass of the resonance. The resonance in question is the Δ-baryon which has a rest mass of $m_0 = 1232\text{MeV}$. The Δ has seven decay modes and all three types of pion can be created in the final state as allowed depending on the flavour, quark content and interaction of the parent particles.

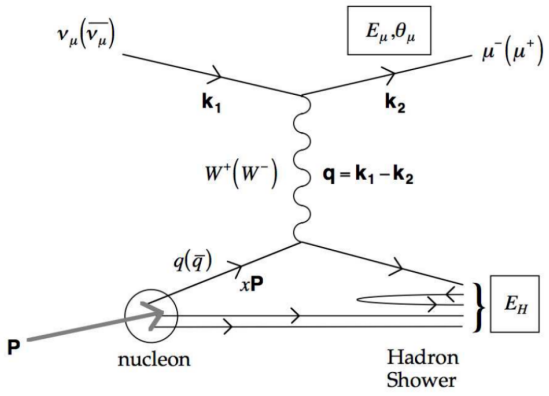


Figure 2.11: Feynmann diagram of the deep inelastic scattering process. Image from [37].

Deep Inelastic Scattering

This process proceeds by the neutrino interacting with one of the constituent quarks of a nucleon, producing a lepton and a hadronic component in the final state. This event is typically catastrophic for the nucleus containing the struck nucleon, and will start a *hadronic cascade*. The process is valid for both neutral- and charged current interactions.

Higher energies

At energies above $\sim$ 10TeV the typical picture of neutrino cross sections change. The propagator term is no longer dominated by the exchange boson mass, which, combined with suppression from a $(1-y^2)$ term leads to an approximate power-law behavior at these energies [37]. One notable exception exist at an energy of 6.3PeVat which a resonance condition is predicted to produce a W-boson. This was first proposed by Glashow in 1960 as a way to directly observe the W-boson and is now known as the *Glashow resonance* [37].

Observation

To observe neutrino via the interactions discussed in this section a specifically designed detector is needed, which is the subject of chapter 3. First, a few more concepts and developments in neutrino physics will be discussed.

2.8 The Solar Neutrino Problem

Nuclear processes in the interior of the Sun produce neutrinos as various energies. The energy of the Sun comes from fusion processes, mainly the 'pp'-chain and and the CNO-cycle. Proton fusion proceeds via the weak interaction in a beta process, yielding neutrinos. The various fusion branches are illustrated in fig. 2.12 and since the energy output of the Sun is known, the expected neutrino flux was calculated by R. L. Sears in 1964 [20]. Around the same time propositions for detection of

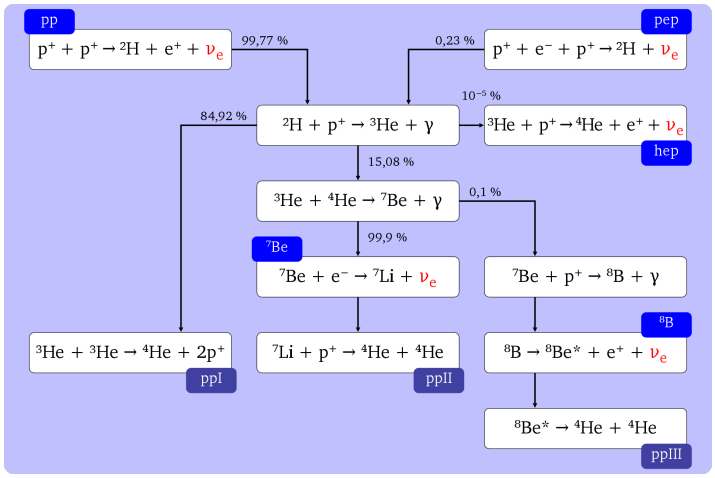


Figure 2.12: Proton fusion processes in the Sun. Neutrinos generated are marked in red. Figure from [28]

these solar neutrinos were put forth by R. Davis and John Bachall, who suggested neutrino capture on Chlorine-37 [21][22], by the reverse electron capture reaction:

$$\nu_e + {}^{37}\mathrm{Cl} \rightarrow e^- + {}^{37}\mathrm{Ar}, \tag{2.17}$$

resulting in a build-up of the radioactive gas Argon. A large reservoir containing the common cleaning fluid Tetrachloroethylene (C_2Cl_4) was dug into the Homestake mine at a depth of about 1.5 km, to act as a neutrino target. The Argon would slowly build up in the tank over time, meaning it could be periodically filtered and the activity level measured. Further reading on the technical details and results of thirty years of running the Homestake experiment can be found in [23]. The results of the neutrino measurements were surprising: Davis and colleagues expected to measure about 7.2 Solar Neutrino Units (SNU), however, only $2.56 \pm 0.16(\mathrm{stat}) \pm 0.16(\mathrm{syst})$ SNU were observed. The measured event rate was thus only about one third of what the calculations from Sears and Bachall had predicted. Later experiments such as GALLEX [24], Kamiokande [25] and SAGE[26] also found a deficit of neutrinos [29]. Either the standard solar model was incorrect, the experiments were flawed or neutrinos were somehow disappearing. This discrepancy became known as the Solar Neutrino Problem.

Although based on Kaon mixing [33], Bruno Pontecorvo is generally attributed as the scientist who came up with a proposed solution to the Solar Neutrino Problem with his suggestion of neutrino oscillations [31, 32]. He proposed neutrinos can change from one flavour to another during propagation. The Sudbury Neutrino Experiment (SNO) later found compelling evidence for solar neutrino oscillations [30]. The detector is located in the Creighton mine in Canada and is currently know as SNO+. The original design contained a 6m diameter acrylic vessel containing about one kTon heavy water (D_2O), and was sensitive to all neutrino flavours via the following interactions [30]:

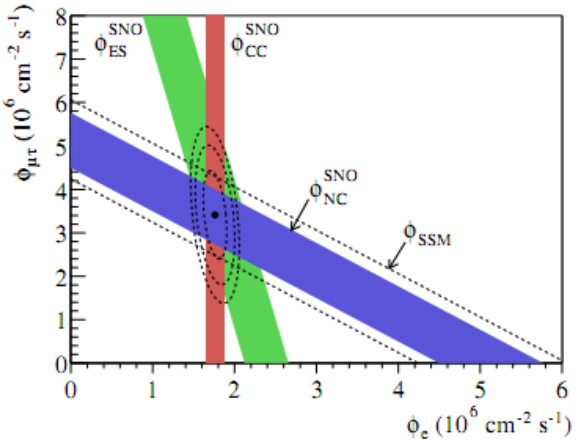


Figure 2.13: Solar Neutrino results from the SNO experiment. The colored bands indicate the fluxes measured with the neutral current reaction (blue), charged current reaction (red) and elastic scattering (green) . The black point shows the best-fit point indicating a combined flux of muon and tau neutrinos to be about twice that of electron neutrinos. Image from [30].

$$A)\nu_e + D \rightarrow p^+ + p^+ + e^- \quad \text{(cc)} \tag{2.18}$$

$$B)\nu_x + D \rightarrow \nu_x + p^+ + n^0 \quad \text{(nc)} \tag{2.19}$$

$$C)\nu_x + e^- \rightarrow \nu_x + e^- \quad \text{(es)} \tag{2.20}$$

Here ν_e refers to the electron neutrino and ν_x refers to any active neutrino flavor. Since reaction A is valid only for electron neutrinos, while reactions B and C are valid for all neutrino flavors, SNO was able to fit for separate flavour content in their data analysis. Their main result for the flavor fluxes is shown in fig. 2.13, and clearly shows the combined muon neutrino and tau neutrino fluxes to be about a factor of two larger than the electron neutrino flux [30]. This was smoking gun evidence for the predicted neutrino oscillations, but more evidence was needed before the puzzle of disappearing neutrinos could be said to be fully solved. This evidence came from measurement of atmospherically produced neutrinos, most notably from the Super Kamiokande experiment [34].

2.9 Atmospheric Neutrino Problem

As discussed in section 2.5 neutrinos are generated by cosmic ray interactions in the atmosphere. The ratio of muon neutrinos to electron neutrinos is expected to be around 2 (at ~GeV scale energies and below). While the early experiments in the 1960s were able to detect atmospheric neutrinos, they were unable to detect their direction, and were lacking in statistics for more detailed investigations. Later experiments, originally intended to search for proton decay, were also sensitive to atmospheric neutrinos and their directionality. In particular the Kamiokande detector and its successor Super-Kamiokande made headway in this area [41, 34]. Based on detailed calculations from M. Honda et al [40]. the ratio of muon neutrinos to electron neutrinos was expected to be constant over the entire

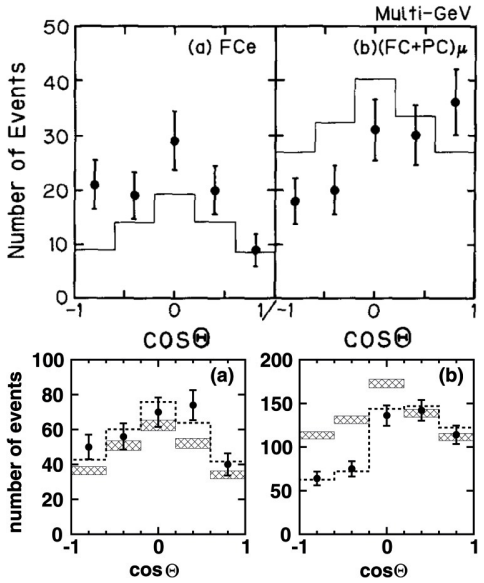


Figure 2.14: Evidence from Kamiokande (top) and Super-Kamiokande (bottom) for atmospheric neutrino asymmetry between up and downgoing regions. Plots show number of neutrino events as function of zenith angle. Black dots with error bars are data. The left-panel shows electron-like events, while the right panel shows muon-like events. Top: The solid line shows prediction based on the HKKM model. Bottom: The boxes show expectation from the HKKM model and the dashed line shows a fit for neutrino oscillations. Figures from [41][34]

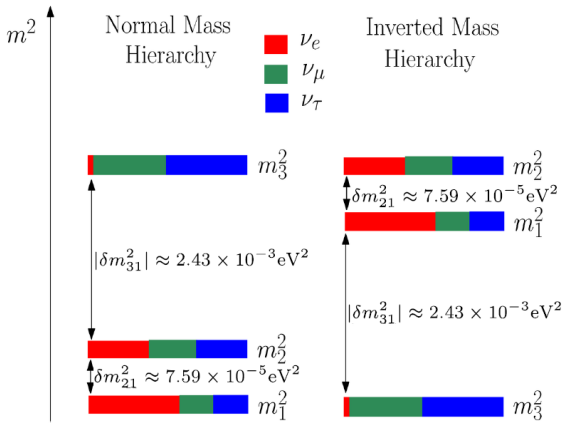


Figure 2.15: Neutrino mass eigen states, their flavour composition and square mass differences. Figure from [43]

sky. However, this was not what Kamiokande observed: as illustrated in fig. 2.14 the experiment saw a clear asymmetry between the up and downgoing regions [41]. Also shown is a later measurement by Super-Kamiokande, with improved statistics, showing clear evidence of neutrino oscillations. A brief discussion of neutrino oscillations will follow in the next two sections.

2.10 Neutrino Oscillations and Mass

As discussed, by the mid 1990's several experiments such as SNO, Kamiokande and Super-K had observed discrepancies in the neutrino event rates when compared to theoretical predictions. Despite Bruno Pontecorvos' early suggestions of neutrino oscillations around 1960 [31, 32], the idea did not gain much traction because mixing angles were a-priori believed to be small, since this is the case in the quark sector [42]. However, with the compelling evidence from several experiments, the idea started to gain prominence.

During propagation neutrinos can undergo a flavor change. The neutrino mass eigenstates can be expressed as super positions of the flavor eigenstates as depicted in fig. [43]. The process is described by the PMNS-matrix in analogy to mixing in the quark sector :

$$U = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} \tag{2.21}$$

The matrix is comprised of three mixing angles and three phases, and can be parametrized in the standard picture as follows:

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \cdot \begin{bmatrix} c_{13} & 0 & s_{13}e^{i\delta_{\text{cp}}} \\ 0 & 0 & 0 \\ -s_{13}e^{i\delta_{\text{cp}}} & 0 & c_{13} \end{bmatrix} \cdot \begin{bmatrix} c_{21} & s_{21} & 0 \\ -s_{21} & c_{21} & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e^{i\eta_1} & 0 & 0 \\ 0 & e^{i\eta_2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{2.22}$$

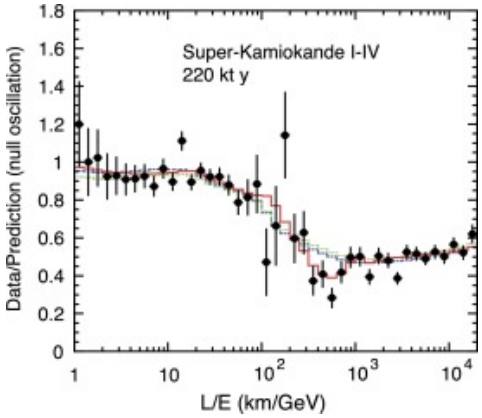


Figure 2.16: Fit for L/E dependence of neutrino oscillation from Super Kamiokande. [34]

with $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$. The mixing angles $(\theta_{12}, \theta_{23}, \theta_{13})$ control the level of mixing between the states while the phases $(\delta_{\text{cp}}, \eta_1, \eta_2)$ control cp-violation. The parallel to ordinary rotation is evident from the shape of the matrices involved. The flavour states can be expressed as a linear super position of the mass states via the transformation:

$$|\nu_\alpha\rangle = \sum_i U^*_{\alpha i} |\nu_i\rangle \tag{2.23}$$

In the standard picture the PMNS-matrix is thought to be unitary, but it can be expanded to accommodate inclusion of oscillation into other theoretical flavours such as one or more sterile neutrinos. In this work, focus will be on the standard picture of neutrino oscillations.

2.10.1 Oscillation Probability

The probability for a neutrino to change flavour during propagation is called the oscillation probability (P), with the non-oscillation case (1-P) called the survival probability. Here, a brief discussion of the probability calculation will be given. Considering the neutrinos as plane waves the probabilities can be found from the time evolution of the mass states as given by the Schrödinger equation:

$$i\frac{d}{dt}|\nu_i\rangle = E_i|\nu_i\rangle \tag{2.24}$$

where E is the energy of the neutrino. The solution can be written as:

$$|\nu_i(t)\rangle = e^{-iE_i t}|\nu_i(0)\rangle \tag{2.25}$$

The mass states oscillate at different frequencies, depending on their energy. Plugging in the flavor/mass state relation of eq. 2.23:

$$|\nu_\alpha(t)\rangle = \sum_i U^*_{\alpha i} e^{-iE_i t}|\nu_i\rangle, \tag{2.26}$$

it is now possible to construct the final transition probability as:

$$P(\nu_\alpha \to \nu_\beta) = |\ \langle \nu_\beta(t) | \nu_\alpha \rangle\ |^2 = |\sum_i U_{\beta i} U^*_{\alpha i} e^{-iE_i t}|^2 \tag{2.27}$$

Of particular interest is the energy term in these equations. From special relativity the energy is given as:

$$E_i = \sqrt{p_i^2 + m_i^2} \tag{2.28}$$

The absolute mass of the neutrino is unknown, but both cosmology and the beta decay spectrum constrain it to be small. Thus in the ultra relativistic limit where $p \gg m$ the energy can be approximated by a Taylor expansion:

$$E_i \approx E + \frac{m_i^2}{2E} \tag{2.29}$$

where E represents the total energy of the particle. Getting to the oscillation probability requires some arithmetic: first rewriting the exponential term; remembering Euler's formulae; the trigonometric double angle relations; and that $t = L$ in natural units:

$$|e^{-iE_i t}|^2 = e^{i(E_i - E_j)t} = e^{i(E + \frac{m_i^2}{2E} - E + \frac{m_j^2}{2E})} = e^{i(\Delta m^2_{ij}/2E)t} \tag{2.30}$$

$$= 1 - 2\sin^2\left(\frac{\Delta m^2_{ji} L}{4E}\right) + i \sin\left(\frac{\Delta m^2_{ji} L}{2E}\right) \tag{2.31}$$

Expanding the matrix part of eq. 2.27 and plugging in the exponential term, while remembering that in the ultra relativistic limit $v_\nu = c$, the oscillation probability takes the form:

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4 \sum_{i<j} \Re\left(\sum_i U^*_{\beta i} U_{\alpha i} U_{\beta j} U^*_{\alpha j}\right) \sin^2\left(\frac{\Delta m^2_{ji} L}{4E}\right) \tag{2.32}$$

$$+ 2 \sum_{i<j} \Im\left(\sum_i U^*_{\beta i} U_{\alpha i} U_{\beta j} U^*_{\alpha j}\right) \sin\left(\frac{\Delta m^2_{ji} L}{2E}\right) \tag{2.33}$$

The mixing angles are then introduced when substituting in the mixing matrix from eq. 2.22. In order to calculate the oscillation probability for anti-neutrinos instead of neutrinos, make the transformation: $U \to U^*$, which will result in a sign change on the imaginary part of the oscillation probability. Note how the phase of the sine function depends both on the square mass difference of the oscillating neutrinos, and also on the $\frac{L}{E}$ fraction. This implies experiments to test various hypotheses regarding Δm^2 must be carefully designed with respect to propagation distance and neutrino energy in order to optimize the statistical signal of the measurement. For the purpose of performing numerical calculations of the neutrino oscillation probabilities the prob3 neutrino oscillation package is used [44].

2.11 Neutrino Propagation in Matter

The calculations in the previous section hold true for neutrino propagation in vacuum. In matter, the picture changes somewhat due to coherent forward scattering. While propagating in a medium, the

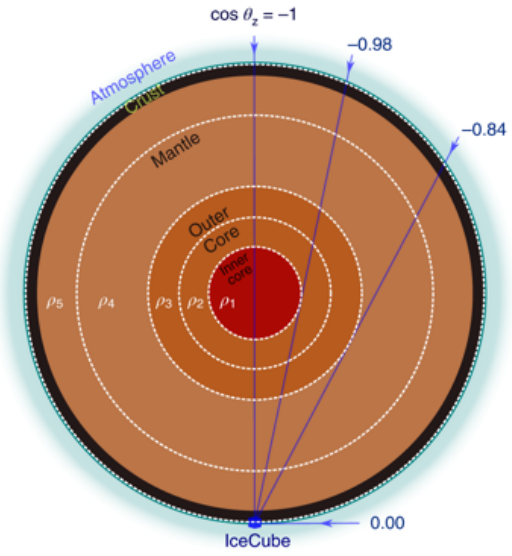


Figure 2.17: Schematic of the Earth layer composition and atmospheric neutrino propagation paths towards IceCube. This illustrates the variation in oscillation baseline as depending on arrival angle. Figure from [48].

neutrinos interact with said medium via the interactions described in section 2.7. When propagating through electron dense environments such as the interior of the Sun or the Earth's core, electron (anti) neutrinos are prone to these effects. All neutrino flavors undergo interactions with electrons via the weak interaction Z^0 exchange. However, the charged current interaction with electron via $W^{\pm}$ exchange is forbidden for ν_μ and ν_τ (and their anti-neutrinos) due to flavor conservation. The effect of these forward scatterings can drastically change the oscillation probabilities for the neutrinos - even in the case of small mixing angles and when ν_μ and ν_τ only undergo neutral current interactions. The effect is known as the Mikheyev-Smirnov-Wolfenstein (MSW) effect, and is not generally restricted to ν_e [47]. In an environment with a high density of the SU_2 counterparts to the ν_μ and ν_τ they would also undergo charged current interactions in a similar fashion. For the electron neutrinos propagating through the earth, the MSW-effect gives rise to them experiencing an extra potential of $V = V^{\mathrm{cc}} + V^{\mathrm{nc}}$:

$$V^{\mathrm{cc}}_{\nu_e} = \sqrt{2} G_F N_e \tag{2.34}$$

$$V^{\mathrm{nc}} = -\frac{G_F}{\sqrt{2}} N_e \tag{2.35}$$

where G_f is the Fermi coupling constant and N_e is the electron number density. As anti-neutrinos have opposite weak Isospin from neutrinos the sign of the potential is inverted in their case: $V_{\bar{\nu}_e} = -V_{\nu_e}$. Since the potential changes the Hamiltonian, the matter Hamiltonian must be written with the extra term:

$$H_{\mathrm{M}} = H_0 + V_{\nu_e} \tag{2.36}$$

This rise in potential is equivalent to an increase in the effective mass of the neutrino. Since neutrino oscillations are not sensitive to the absolute scale of the mass but only to the square mass difference, the neutral current contribution, which is valid for all neutrino flavors, can be ignored. The Hamiltonian in the two-flavor approximation can then be shown to take the following form [49]:

$$H_{\mathrm{M}} = \left(\frac{\Delta m^2_m}{4E}\right) \begin{bmatrix} -\cos(2\theta) + A & \sin(2\theta) \\ \sin(2\theta) & \cos(2\theta) - A \end{bmatrix} \tag{2.37}$$

where the factor A is given as:

$$A = \frac{2\sqrt{2} G_F N_e E}{\Delta m^2} \tag{2.38}$$

With this it is then straightforward to recalculate the Schrödinger equation from eq. 2.26, in terms of the effective square mass difference in matter Δm^2_{M} and the effective mixing angle in matter θ_{M}. It then takes the familiar form:

$$i\frac{d}{dt}\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = \frac{\Delta m^2_{\mathrm{M}}}{4E} \begin{bmatrix} -\cos(2\theta_{\mathrm{M}}) & \sin(2\theta_{\mathrm{M}}) \\ \sin(2\theta_M) & \cos(2\theta_{\mathrm{M}}) \end{bmatrix} \begin{pmatrix} \nu_i \\ \nu_j \end{pmatrix} \tag{2.39}$$

Introducing the constant C:

$$C^2 = (\cos(2\theta) - A)^2 + \sin^2(2\theta), \tag{2.40}$$

the effective parameters can be shown to take on the following values:

$$\Delta m^2_{\mathrm{M}} = C\Delta m^2 \tag{2.41}$$

$$\sin(2\theta_{\mathrm{M}}) = \frac{\sin(2\theta)}{C} \tag{2.42}$$

A few comments are in order:

1. The extra potential gives rise to a resonance under the condition $\cos(2\theta) = A$, meaning oscillations can be enhanced irrespective of the magnitude of the mixing angle θ.
2. The resonance occurs only for $A > 0$ ($A < 0$ for anti neutrinos), and since A depends on Δm^2 this can be used to determine the sign of the square mass difference, and solve the neutrino mass hierarchy. If, as is the case for solar neutrinos, the resonance occurs, then Δm^2 must be manifestly positive.
3. A scales with electron number density and energy, and at low values of A, the vacuum oscillation picture is recovered. For the MSW effect to be pronounced, high electron density or long baselines are required. (while A scales with energy, the phase of the oscillation is inversely proportional to the Energy, effectively canceling out the enhancing effect.)

For neutrino detectors such as IceCube, the MSW effect plays an important role. Atmospheric neutrinos from cosmic ray interactions travel from their interaction point, and since they happen all over the planet the propagation distance will vary from about 20km up to the diameter of the Earth. The baselines are exemplified in fig. 2.17. The impact from oscillations on IceCube measurements are then the topic of the next section.

2.12 Oscillation impact on an IceCube Neutrino Measurement

Neutrino oscillations will impact the measured spectrum in both energy and arrival angle and will have a profound effect on the location of the oscillation minimum. This analysis uses the oscillation parameters as input, thus the uncertainty related to them must be considered. As described in section 2.10, neutrinos undergo oscillations, the magnitude of which depend on the energy and propagation length. The neutrinos measured in this analysis fall in the energy range from 1GeV to 1TeV and propagate baselines of 20km (directly downgoing) to 12700km (directly upgoing), making their measurement very sensitive to neutrino oscillations. The critical factor for oscillations is L/E, which in this case falls in the range: $0.02\frac{\text{km}}{\text{GeV}} \leq L/E \leq 635\frac{\text{km}}{\text{GeV}}$. At these values the oscillation effect from the first mass difference $\Delta m^2_{21} = 7.56 \cdot 10^{-5} eV^2$ is expected to yield only a small impact on the event rates (as the effect does not reach the percent level until L/E reaches about $10^3\frac{\text{km}}{\text{GeV}}$). However, we can expect an impact from the second mass splitting at $\Delta m^2_{31} = 2.55 \cdot 10^{-3} eV^2$ with a mixing angle of $\theta_{23} = 41°$, because the L/E range corresponds to about one period of oscillations, as illustrated in fig. 2.18.

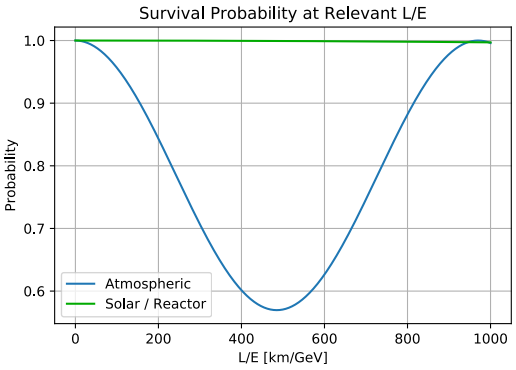


Figure 2.18: Survival probability of reactor- and atmospheric neutrinos as a function of L/E. Notice the solar/reactor survival probability drops by less than 1% over the relevant L/E range.

The cp-violating phase will impact the oscillations differently for neutrinos and anti-neutrinos yielding slightly different neutrino to anti-neutrino ratios, and the cp asymmetry can be expressed as:

$$P\left(\nu_\alpha \to \nu_\beta\right) - P\left(\bar{\nu}_\alpha \to \bar{\nu}_\beta\right) = 4 \sum_{i>j} \left(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\right) \sin\left(\frac{\Delta m^2_{ij} L}{E}\right) \tag{2.43}$$

This is thus manifested as an asymmetry in the oscillation pattern between neutrinos and anti-neutrinos. While Δm^2_{ij} is expected to stay the same in both cases due to symmetry reasons, the cp-violating phase will have an impact on the matrix term. Furthermore, the neutrino interaction cross sections are different between neutrinos and anti-neutrinos leading to an enhancement of these asymmetries at final level. The oscillations are calculated using the Prob3 calculation code [44], including three flavours, matter effects and including a cp-violating phase, and the oscillograms for the simulation used in this dissertation are shown in fig. 2.19.

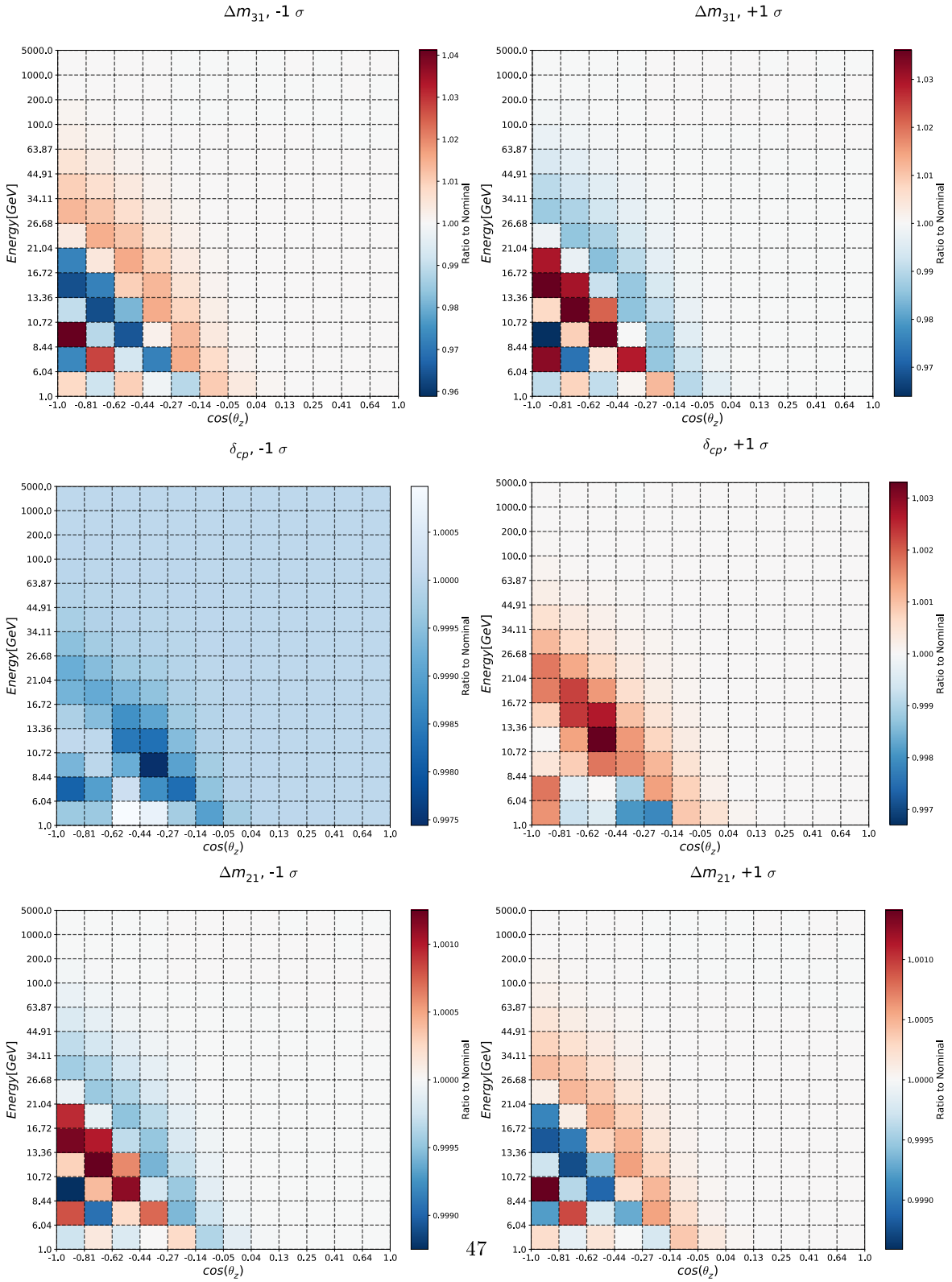


Figure 2.19: Oscillation systematics impact on the simulated spectrum. Notice here that although the colors for Δm_{21} and δ_{cp} might look dramatic the scale is down to the per-mill level. It is thus reasonable to expect negligible impact from these systematics in the analysis.

Oscillation Parameters		
Parameter	Value	Prior
θ_{12}	$34.5°$	$\pm 1.1°$
θ_{23}	$41°$	$\pm 0.11°$
θ_{13}	$8.41°$	$\pm 0.17°$
Δm_{21}	$7.56 \cdot 10^{-5} eV^2$	$0.19 \cdot 10^{-5} eV^2$
Δm_{31}	$2.55t \cdot 10^{-3} eV^2$	$0.04 \cdot 10^{-3} eV^2$
δ_{cp}	$252°$	$\pm 24°$

Table 2.3: Table of neutrino oscillation parameters used in this analysis, taken from global fit in [50].

3 The IceCube South Pole Neutrino Observatory

IceCube is a neutrino telescope located at the geographic south pole. This chapter is mainly concerned with a brief technical description of the telescope, its components and technical capabilities. A more in-depth treatment can be found in [51]. In short, 5160 digital optical modules, called DOMs, have been deployed in the deep layers of ice at depths between 1450m and 2820m [52]. The DOMs are connected on kilometer long cables to the IceCube laboratory at the Antarctic surface. The cables are colloquially known as strings and IceCube is comprised of 86 of these. 60 DOMs are hosted on each string, which are organized in a hexagonal pattern. The typical string spacing is 125m and the inter-DOM spacing on the strings is 17m, making the instrumented region 1km high and 1km across. At the center of the detector is the "DeepCore" sub-array specifically designed for detection of neutrinos with energies in the tens of GeV range. A surface array of Cherenkov detectors, called IceTop, is utilized to detect cosmic ray induced air showers, and can also be used as a veto for the in-ice detector array. A sketch of the detector layout and geometry is shown in fig. 3.1.

The Antarctic ice at these depths are under immense pressure, greatly compacting the ice and forcing out impurities to the point that the ice is the clearest and most transparent non-gaseous material on the planet. In the deeper regions of the detector, the average absorption distance for a photon of 400 nm wavelength goes all the way up to 264m with typical values around 180m [74]. A notable exception to this is the presence of a 'dust layer': a region of ice with high absorption and scattering due to foreign bodies in the ice, likely soot and dust from volcanic activity during the corresponding geological time period.

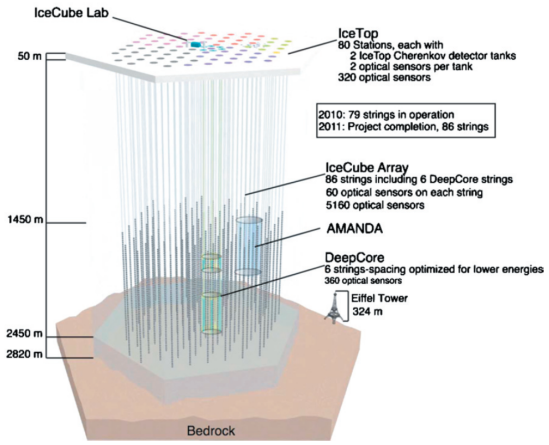


Figure 3.1: IceCube detector architecture schematic, with the Eiffel tower for scale reference.

3.1 DeepCore

In the bottom center of IceCube, 8 strings have been organized with a denser and asymmetric string spacing. This region is known as DeepCore and has been designed for GeV-scale measurements. This denser instrumentation allows for a lower energy threshold and better angular and energy resolution for events in this volume. The DeepCore instrumentation itself consists of 60 DOMs per string, like the

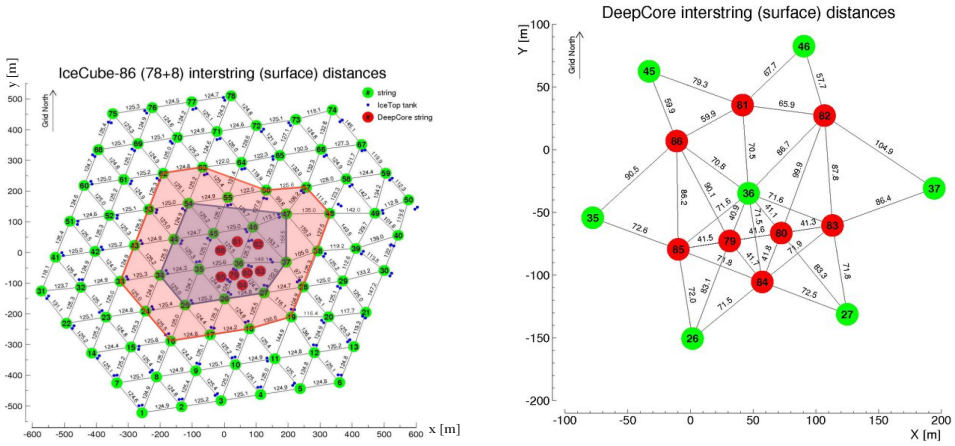


Figure 3.2: Layout of the IceCube/DeepCore string geometry. Green circles represent ordinary IceCube strings, with red circles marking DeepCore strings. The shaded regions indicate the surface area of the trigger volume for the standard DeepCore configuration (purple) and the 2-Layer veto (pink). Figures from [54]

ordinary IceCube strings, however, the vertical distance between the modules is only 7m as compared to the 17m of the standard IceCube configuration. The geometry is shown in fig. 3.2. Moreover, the DeepCore DOMs are fitted with PMTs, which have a higher quantum efficiency than the standard IceCube DOMs. This combines to make the DeepCore region more light sensitive than the rest of IceCube, effectively lowering the energy threshold for neutrino detection to around 10 GeV [52]. The DOMs included in the DeepCore trigger are all the DOMs on the 8 DeepCore strings, as well as the central string 36, and the inner layers of ordinary IceCube strings. Data taken from DeepCore is the main input for the analysis in this work.

3.2 Detection Principle: Cherenkov Radiation

A charged particle propagating faster than the phase velocity in a dielectric medium will emit Cherenkov radiation along its path. Such charged particles can be created by neutrino interactions in the ice, and the light they emit can be detected via photo sensors embedded in the ice.

The Cherenkov light is emitted by the medium polarized by the traversing particle - when the velocity of the traversing particle is sufficiently large, ie faster than the phase velocity of light, the polarization becomes asymmetric and radiation is emitted, as illustrated in fig. 3.3. The emitted radiation forms a wavefront of constructive interference, giving rise to a characteristic cone shape emission as shown in fig. 3.4. This is highly analogous to the waves generated by a boat moving faster than the phase velocity of water waves, or to the sonic boom created by an aircraft traveling at super sonic speed. The opening angle of the emission cone is well known and defined by the index of refraction n and the velocity of the particle v as related to the speed of light c:

$$\cos(\theta) = \frac{c}{vn} \tag{3.1}$$

For ultra relativistic particles the opening angle depends only on the index of refraction and eq.

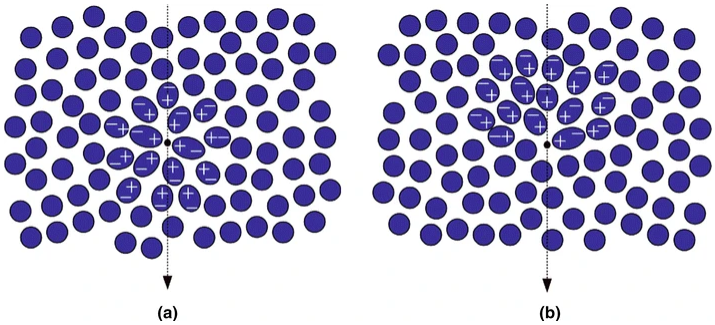


Figure 3.3: Polarization of a dielectric medium by charged particles of different velocities. Left: Velocity is smaller than $\frac{c}{n}$ and the polarization is symmetric. Right: Velocity is greater than $\frac{c}{n}$, the polarization becomes asymmetric and radiation is emitted . Figure from [55]

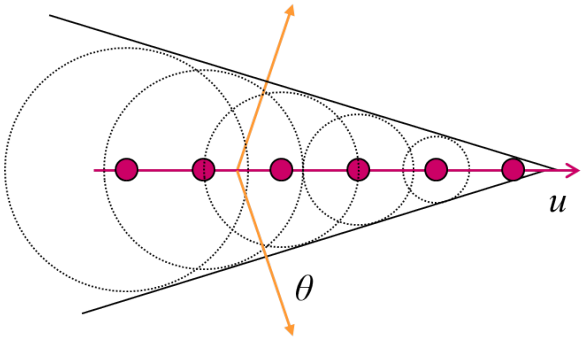


Figure 3.4: Geometry of Cherenkov radiation emission from a dielectric medium following penetration of a charged particle of velocity u. Light waves are emitted spherically and a characteristic cone shape shock front is created by constructive interference. Figure from [56]

3.1 can be used in event reconstruction to estimate flight times of photons incident on the optical modules. The main detection principle thus relies on emission of Cherenkov radiation caused by secondary particles from neutrino interactions, and their collection on the optical modules in the ice.

3.3 Photo Sensors

The photo sensors employed in IceCube are classical 10 inch *photomultiplier tubes* (PMT), housed in individual containment units called *Digital Optical Modules* (DOM) [51]. Each DOM is packed with a downward facing PMT as well as on board circuitry for immediate data processing, including data acquisition, calibrations, and communication with the surface. The whole unit is packed in a pressure resistant vessel, a schematic is shown in fig. 3.5. PMTs are capable of taking in photons and giving off an electric signal. A photo cathode is located in the front of a glass vacuum tube, as a photon strikes the photo cathode one or more electrons are emitted through the photo electric effect. Via an electric field the emitted electrons are accelerated towards the back of the PMT, where the signal is multiplied

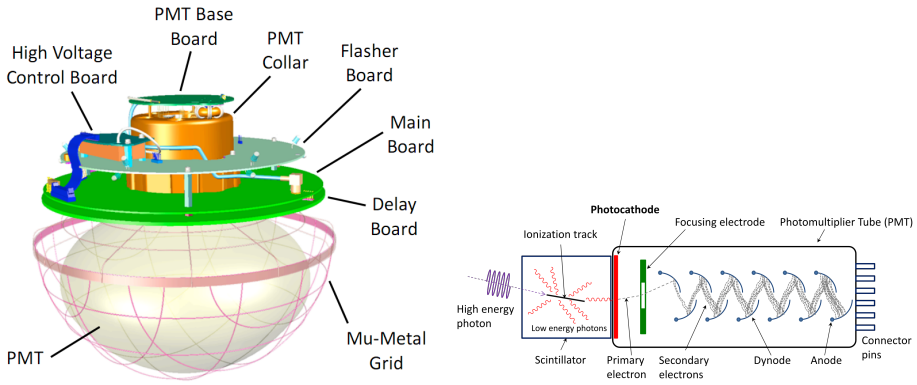


Figure 3.5: Schematic of an IceCube DOM. The DOM is comprised of several components: the downward facing PMT, main electronics board for low level data processing, as well as power supply, and calibration devices. [51]

by about a factor of several million, via a series of cascading dynodes, before hitting the anode in the back. This process facilitates detection of single *photo electrons* (PE). The IceCube PMTs are produced by Hamamatsu and are sensitive to photons in the 300nm - 650nm wavelength range, with a peak quantum efficiency (QE) of about 25% (34% for the high QE DOMs positioned in DeepCore) [51]. The main task of the DOMs is to detect the light emanating from the particle interactions in the ice and digitize the signal.

3.3.1 Waveforms

The analogue electron signal is first registered on the anode where it creates a voltage drop. This is known as a *pulse*. The basic data recording is governed by an on board discriminator set to a pulse level of 0.25 photoelectrons. When the threshold is surpassed, a recording sequence known as a DOM launch is begun, which constitutes the lowest level signal information in IceCube. The signal is digitized by on board dual-channel *analogue transient waveform digitizers* (ATWD), at a sampling rate of 300 Msps over a duration 427 ns. Each channel also has three gain configurations of 16, 2, and 0.25 to adequately cover the dynamic range of the PMT output. The 427ns time window is sufficient to record photon signals from within a radius of tens of meters from the DOM. Photons from farther away may display delayed timing properties due to scattering in the ice, and will also be lower in amplitude. To capture these, the DOMs are fitted with *fast analogue digital converters* (fADC) continually sampling the waveform over a time window of 6.4 μs. Some deadtime is expected as the ATWDs have a ramp-up time of about 29μs. Examples of digitized waveforms are shown in fig. 3.6. Digitized waveforms are reconstructed utilizing the WaveDeForm module, yielding a best-fit estimate of the waveform prior to digitization called a *reconstructed pulse*. Launches on individual DOMs are not sufficient to categorize an event, thus the physics data acquisition system operates with coincidences, which is the topic of the next section.

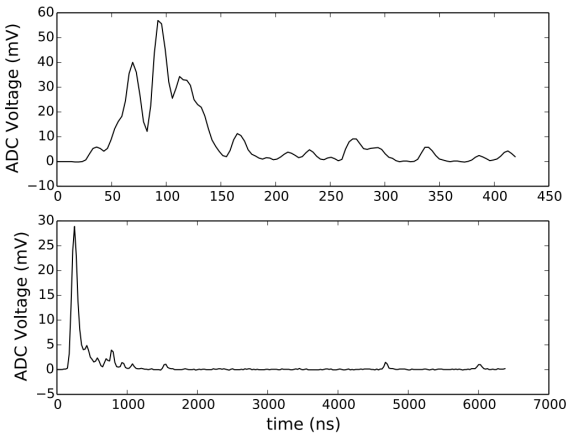


Figure 3.6: Example of a Digitized waveform from an IceCube DOM. The sample signal is sampled by an ATWD (top) and fADC (bottom) over 427ns and 6.4μs respectively. [51]

3.3.2 Trigger Conditions

When a DOM launches, the nearby DOMs are notified of that launch via hardware coincidence wiring. Any coincident launches on the two DOMs above or below on the same string within $\pm 1\mu s$ are marked as *hard linear coincidences* abbreviated *HLC*. Launches inside the time window, which fail to satisfy these conditions are categorized as soft linear coincidence (SLC). The number of HCL-hits and SLC-hits are used to define software trigger conditions in the physics data acquisition system (pDAQ). The most commonly used trigger type is the *simple majority trigger* (SMT), which searches for HLC hits in space and time to group DOM launches into into collections constituting events. The trigger conditions are based on three parameters: A time window, a set of DOMs, and the number of hits required. Several SMT trigger algorithms run in parallel with different parameter settings. The data used for the analysis presented in this dissertation is based on the *DeepCore SMT3 trigger:* it uses the DeepCore DOMs plus the first layer of surrounding IceCube strings, requires 3 HLC hits, and has a time window of 2.5μs. The pulse information of all hits in an event is stored in a data object called a *pulse series*, which is used for later processing and data selection. The relatively wide time window of the trigger can lead to clusters of DOMs being included in an event from quite separate locations in the detector, necessitating further cuts on the data. This is discussed in section 4.2, which covers the sample selection.

3.4 Effective Area

The effective area represents the probability for a particle to interact within the fiducial volume of the detector. The name comes from the geometric considerations regarding cross section, but is modified due to the fact that incoming particles interact not based on their geometric size, but due to a fundamental force interaction. Including the factors based on that ansatz, the 'effective area' is a measure of the cross section IceCube effectively represents to an incoming particle. The effective area is calculated from the interaction cross section of the fundamental interaction in question, via the Feynman rules, multiplied by the number of targets (in the non-shadowing case), and is a measure of

the probability for particles of interest to interact in the ice (in this case, neutrinos).

- Number Density can be calculated from H_2O molar mass: $M_{\text{mol}} = 18.01528\frac{g}{\text{mol}}$, Avogadro's number $N_A = 6.22 \times 10^{23}/mol$, the density of the ice: $\rho = 0.9167\frac{g}{cm^3}$ and the fiducial volume of IceCube/DeepCore. Under ideal conditions, the ice consists solely of H_2O molecules in the ice phase, and is homogeneous and isotropic. Any deviation from this assumption will be treated as a systematic uncertainty in section 7.

$$M_{mol} = 18.01528 g/mol \tag{3.2}$$

$$\rho = 0.9167 \left[\frac{g}{cm^3}\right] \tag{3.3}$$

$$n = \frac{\rho}{M_{mol}} N_A \tag{3.4}$$

$$= 6.22 \times 10^{23} \frac{1}{mol} \frac{0.9167\frac{g}{cm^3}}{18.01528\frac{g}{mol}} = 4.7455 \cdot 10^{22} cm^{-3} \tag{3.5}$$

- n is then the number of water molecules per cubic centimeter of ice. Each Oxygen nucleus consists of 16 nucleons, with the two Hydrogen atoms each contributing one proton to the molecule, giving a total of 18 nucleons: 10 protons and 8 neutrons.

- The total effective area is then a sum over all the different neutrino flavours and their interaction types, multiplied by the number of targets in the IceCube fiducial volume. The interaction types have been described in further detail in section 2.7.

$$A_{\text{eff}} = V_{\text{fidu}} \sum_{\text{flav}} \sum_{\text{int}} \sigma_{\text{int}}^{\text{flav}} n_{\text{target}}^{\text{flav,int}} \tag{3.6}$$

$$\left(\sigma_{cc}^{\nu_\mu} + \sigma_{cc}^{\bar{\nu}_\mu}\right) N_{target}^{cc} \tag{3.7}$$

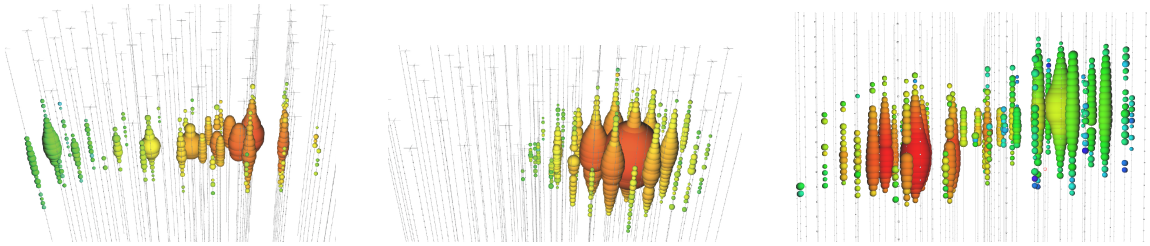
The effective area then correctly has units of area $\left[\text{cm}^2\right]$. The number of targets vary by interaction type, as some combinations are only valid for certain interactions such as scattering on electrons or anti-neutrinos incident on protons.

3.5 Event Topology

In an idealized picture IceCube/DeepCore event topologies can be sorted into three distinct categories. They are classified as: track, cascade, composite and double bang, and are signatures of specific particle and interaction types in the detector. As described in section 3.2, charged particles moving faster than the phase velocity of light in the medium emit Cherenkov radiation. Different combinations of particle and interaction type will leave different signatures and event topologies in the detector. This facilitates some degree of particle identification via the various topologies.

Track

Neutrino interactions yielding an outgoing charged lepton, leads to continuous light emission along the trajectory of the charged lepton. Insofar as the interaction vertex is outside the IceCube/DeepCore fiducial volume, this gives rise to a pure track event. Track events are associated with muons, either



(a) Track (b) Cascade (c) Double Bang

Figure 3.7: Typical event topologies in IceCube. (a) Through going track event, associated with muons and ν_μ. (b) Cascade event associated with neutral current interactions and charged current interactions from ν_e and ν_τ. (c) Double bang signature associated with high energy ν_τ charged current interactions.

from cosmic ray induced air showers or from muon neutrino interactions in the ice. TeV-scale track events have a pointing resolution on the sky of about 1 degree, however, since the interaction vertex is outside the fiducial volume, the energy reconstruction of the incoming neutrino becomes more uncertain.

Cascade

Neutral current interactions do not yield an outgoing charged lepton, so the energy of the interaction is deposited into a hadronic or electromagnetic cascade. In both cases the emitted secondaries lead to emission of Cherenkov radiation from a small prolate region in the ice. The region is typically a few meters to a few tens of meters in size, depending on the energy. Such a structure is far smaller than the distance between the DOMs and is thus very difficult if not impossible to resolve. The cascade emits radiation roughly isotropically, leading to an approximately uniform distribution of hits on the DOMs, and is classified as a cascade topology event. These events are associated with all neutral current events, as well as ν_e and ν_τ charged current events, due to the very short propagation distance of the charged lepton in the ice.

Double Bang

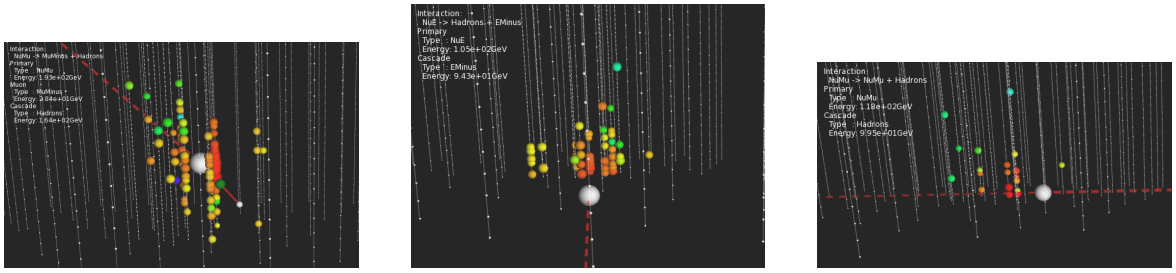
The third type is the double bang, characteristic of a tau neutrino charged current interaction: along with a cascade, a tau lepton is generated at the interaction vertex and propagates outwards. However, due to the short lifetime of the tau it will decay within the fiducial volume giving rise to another cascade. This type of event can only be generated by tau neutrinos of PeV-scale energies and is thus irrelevant for the analysis presented in this work.

Composite

The 'composite' event topology occurs for charged current events taking place within the fiducial volume. They are also called 'starting events', as the initial interaction takes place within the fiducial volume. A cascade is formed at the interaction vertex, and the outgoing charged lepton will give rise to a track signature along its trajectory. This event signature is associated with charged current events from ν_μ and ν_e interactions.

Beyond Ideal

Looking beyond the idealized cases, event topology and particle identification becomes more complicated: at GeV energies very little Cherenkov light is available compared to the DOM spacing, making resolution of the event topology a challenge. A low energy muon neutrino charged current event will send out a Cherenkov muon, however, its energy will be low enough that it will quickly decay leaving only a very short track, effectively mimicking the cascade events. On the other hand, cascade type events can be oriented in such a fashion that the pixelation due to the DOM spacing can mimic a composite event, containing both a cascade and a track. With the low amount of light emitted at GeV energies, the number of DOM hits is limited and the pixelation effect becomes highly pronounced: the idealized event signatures no longer hold and event classification becomes difficult, as shown in fig. 3.8. Effectively, most events at DeepCore energies will be composite and the reconstruction algorithm discussed in the next chapter will always fit for both a track and a cascade component.



(a) ν_μ-cc (b) ν_e-cc (c) ν_μ-nc

Figure 3.8: Examples of non-ideal event topologies taken from MC, where the truth of the incoming particle and its interaction is known. (a) 193 GeV ν_μ charged current event, which in spite of yielding an outgoing charged lepton displays no clear track. (b) 105 GeV ν_e charged current interaction displaying what could be interpreted as a long track. however, at 105 GeV the electron track is expected to be on the scale of 10 cm. (c) 118 GeV ν_e neutral current interaction, with irregular topology in which it is not straight forward to define either a track or a cascade.

3.6 Ice Properties

The IceCube neutrino observatory is built both into and on top of the Antarctic ice sheet. This 2.8 km thick glacier provides both the support structure and the interaction medium for the IceCube observatory. The ice is formed over millions of years from gradual accumulation of precipitation over the seasons. Atmospheric properties over geological time scales are imprinted in the ice through air bubbles and mineral deposits from the atmosphere. An important distinction must be made between the bulk of the deep glacial ice and the newly re-frozen ice in the bore holes. In the following a description of both classifications will be given.

3.6.1 Bulk Ice

The main bulk of the ice is aptly named 'bulk ice' and is made up of the old ice deep in the Antarctic glacier. Three physical parameters are of interest: the optical index of refraction the ice, the light scattering, and light absorption in the ice. Firstly, the index of refraction is related to the emission direction of Cherenkov radiation, and while it is wavelength dependent, it is about 1.38 at the Cherenkov

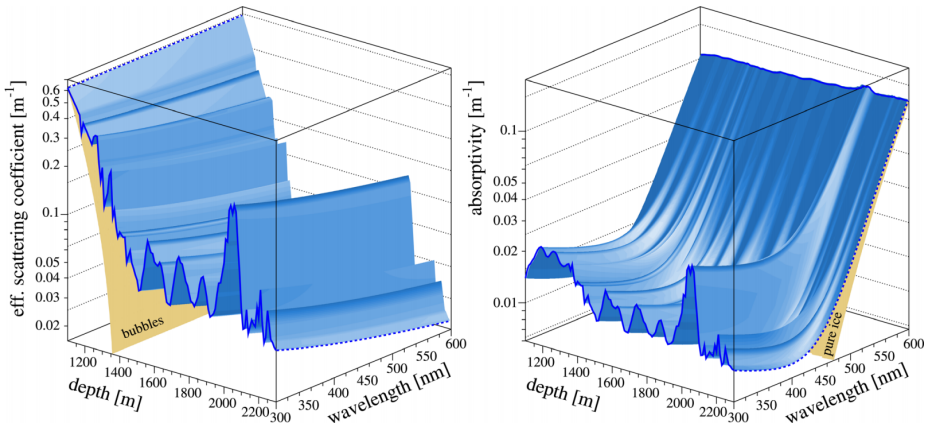


Figure 3.9: Light scattering and absorption in the deep Antarctic glacial ice as a function of depth and wavelength. A region of high scattering and absorption is visible around 2000 m depth, and only has a slight dependence on wavelength.

emission peak of about 325nm [73]. Good knowledge of the emitted wavelength and angle of Cherenkov light facilitates good reconstruction of events. Secondly, light scattering in the ice is governed by photon scattering on electrons, predominantly involving negligible energy transfer through Raylaigh and Mie scattering. Scattering generally has the effect of diffusing a concentrated light signal. In the case of IceCube this will change the shape of emitted light and, combined with the geometry of the detector, will migrate event topologies toward a more cascade-like signature. Thirdly, light absorption is described as the characteristic average length a photon will travel in the medium before being absorbed. Longer distances are preferred for the experimental design, as shorter distances quench light emission from events. Deep in the ice the immense pressure is thought to force bubbles out of the ice, creating a very clear and optically transparent solid. This was first investigated in the AMANDA experiment using both pulsed and continuous light sources embedded in the deep ice [73]. The results are shown in fig. 3.9, and indicate that below depths of about 2.2 km the absorption drops to the 100m scale, with the scattering length being about 20m. A 'dust layer' of high scattering and absorption exists at depths of about 2 km, very likely due to geological periods of high volcanic activity yielding high concentrations of soot in the Earth's atmosphere. These measurements have since been verified by IceCube, and the result from a study done using led flashers on the DOMs, shown in fig. 3.10 gives a similar and consistent picture of the ice properties [74]. $b_e(400)$ is a depth dependent parameter directly related to the scattering of light in the ice. $b_e(400)$ and a corresponding parameter $a_{\text{dust}}(400)$ are well described in [74].

Novel imaging set ups are in production for the IceCube upgrade, in an effort to improve understanding of the deep Antarctic ice.

3.6.2 Hole Ice

The ice in the drill holes displays different physical properties as compared to the main bulk of Antarctic ice. The bore holes housing the IceCube strings are drilled using a high-pressure hot water drill, melting away the ice, and forming a column of liquid water into which the strings are deployed. Over the course

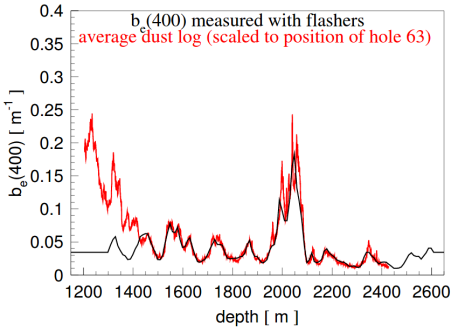


Figure 3.10: Fit to flasher data in IceCube, yielding a measure of the effective scattering coefficient $b_e(400)$

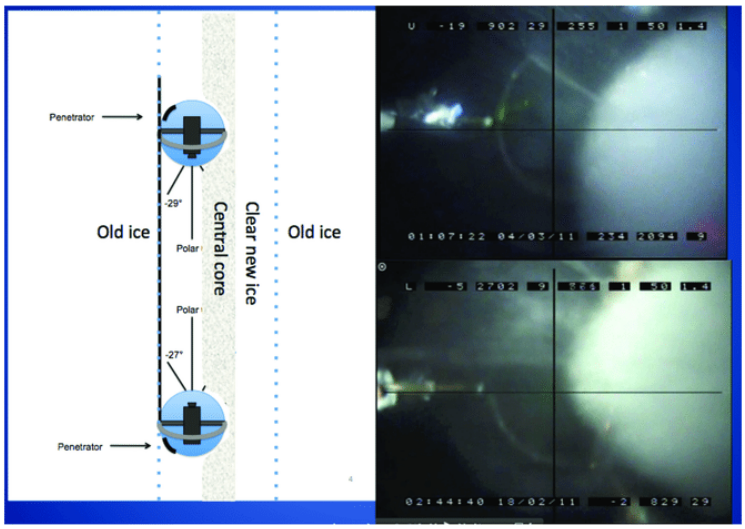


Figure 3.11: Illustrating the bubble column and the formation of hole ice. [53]

of about 10 days the ice then re-freezes encasing the cables and the DOMs in the ice. The ice freezes from the outside towards the center of the column and a curious effect manifests: impurities and air bubbles are pressed out of the freezing ice and into the slightly warmer water towards the center of the column, eventually forming what is known as a 'bubble column', which is a cylindrical volume of increased scattering and absorption of light roughly 10 cm in radius. Depending on the exact radial positioning of the DOM in the bore hole, this obscures part of the main photo cathode area lowering the effective light sensitivity of the DOM. The effect has been studied in several ways: through cameras deployed in the ice [53], in lab experiments, as illustrated in fig. 3.11 and calibrated by using led flashers on the mainboards of the DOMs [51]. The obscuring effect of the bubble column contributes towards the overall systematic uncertainty of the experiment, and its modeling as a systematic parameter will be discussed in section 7.

4 Data Sample and MC Simulation

IceCube collects data continuously save for short periods of downtime. Much of the collected data is unsuitable for the neutrino search reported in this dissertation. As such, the raw data needs to be processed: background events must be rejected, containment cuts enforced, and events reconstructed, before a data sample can be used for analysis. The process is performed sequentially with each individual step known as an *analysis level*. Each analysis level typically performs a specific form of data filtering, utilizing several different variables described below. The entire process from start to finish is simulated in order to optimize the process. This chapter will discuss the data sample and the corresponding simulation used in this work.

4.1 IceCube Low Energy Simulation

A portfolio of standardized neutrino simulation sets exists for IceCube/DeepCore. They have been produced using many different parameters, as a combined effort by many people in the IceCube low-energy working group, with Andrii Terliuk taking the lead on simulation production. Generation of the main neutrino signal was done using GENIE 2.8.6 [57] and drawn from an energy dependent power law with a spectral index of $\gamma = -2.0$. For computational reasons all generated neutrinos are forced to interact. The events are then re-weighted to the expected flavour and neutrino-anti-neutrino ratios. These are based in particular on the expectations from the atmospheric interactions and the cross sections in GENIE. The IceCube/DeepCore detector and decay particles are simulated using GEANT4. Any photons are propagated through the ice using CLsim. Some photons eventually reach the PMTs, potentially trigging the IceCube data acquisition (DAQ) algorithms.

4.2 Data Sample Selection: GRECO

This analysis makes use of the GRECO sample selection, developed by M. Larson for use in several analyses. The sample selection is described in detail in ref. [60], and a short overview will be given here. The event selection consist of seven levels of cuts which proceed from the Level 0 hit cleaning, through various cuts, to the event reconstruction and final selection at Level 7.

4.2.1 Hit Cleaning

In a typical DeepCore neutrino event at about 10 GeV, only about 10-20% of DOM hits are caused by photons stemming from the neutrino interaction. Other hits are caused by *noise hits* happening within the event time window, and contain no information about the physics of the particle interaction. The purpose of the *hit cleaning* step is thus to rid the pulse series of the noise hits and only leave the relevant hits. Several algorithms for hit cleaning exist, and each will generate a new pulse series which is added to the event file. Here, the two hit cleaning algorithms used in GRECO are briefly discussed:

HLC Cleaning

The strictest hit cleaning algorithm in the IceCube software framework. It accepts only DOM launches which satisfy the hard linear coincidence (HLC) conditions as described in section 3.3.2. Thus, the resulting pulse series is cleared of almost all detector noise, but also leaves out a substantial amount of information about the event, as all soft linear coincidence (SLC) hits are excluded. HLC pulse series can be used as input seeds for the *SeededRT* hit cleaning algorithm:

SeededRT Cleaning

As the name implies, this algorithm starts from a seed in the form of a pulse series, usually the series of HLC hits produced by the HLC cleaning algorithm. SeededRT attempts to identify interesting SLC hits based on their proximity to the HLC hits and also include them in the pulse series. For each DOM in the HLC series a sphere of radius r is defined and any SLC hit within it and within a time window Δt will also be included in the pulse series. Since this depends on the detector geometry due to the density of the DOM spacing, the parameters are different for IceCube and DeepCore. IceCube uses $r = 150$m and $\Delta t = 1000$ns while DeepCore uses $r = 75$m and $\Delta t = 500$ns. SeededRT Cleaning allows for high purity pulse series while still locating most hits associated with muon or neutrino events.

Time Window Cleaning

This cleaning algorithm comes in two variants. The first one, *static time window* (STW) cleaning looks for hits around the trigger time. It permits almost all the noise hits to remain in the pulse series. In DeepCore any hits 4μs before the SMT3 trigger or 6μs after the trigger are removed. The second version is a *dynamic time window* (DTW), which finds the time with the highest density of hits and places a window around this time. The DTW is tighter and is typically set to a few hundred nanoseconds, as compared to the 10 μs of the STW. Time window cleaning is typically combined with other cleaning mechanisms. In the case of the DTW, this can result in a high fraction of hits stemming directly from unscattered photons.

4.2.2 L1:DeepCore Filter

The first data filtering algorithm is the DeepCore online filter, which runs continuously at the IceCube Lab at the South Pole. The aim of the filter is to select events within the DeepCore fiducial volume, shown in fig. 3.2 and reject atmospheric muons. The fiducial volume includes a set of DOMs consisting of:

1. The bottom 22 DOMs on strings: 25 - 27, 34 - 37, 44 - 47 and 54
2. The bottom 50 DOMs on strings 79 - 86

The remainder of the DOMs are part of the veto region. The DeepCore filter starts by selecting events containing an SMT3 trigger as described in section 3.3.2. With the pulse series loaded in memory, a Center of Gravity is calculated as the average of the position vector $\vec{r}_i$ to the i'th DOM [52, 60]:

$$\vec{r}_{\mathrm{CoG}} = \sum_{i}^{N_{\mathrm{Hits}}} \frac{\vec{r}_i}{N_{\mathrm{Hits}}} \tag{4.1}$$

The time of the center of gravity can then be calculated:

$$t_{\mathrm{CoG}} = \sum_{i}^{N_{\mathrm{Hits}}} \frac{t_i^0 - \frac{||\vec{r}_i - \vec{r_{\mathrm{CoG}}}||}{c_{\mathrm{ice}}}}{N_{\mathrm{Hits}}} \tag{4.2}$$

with t_i^0 being the time of the first hit in the pulse series. The next logic of the algorithm is based on the effective velocity of the signal from the interacting particle: muons will traverse the detector at close to the speed of light $v \approx 0.3$m/ns, while signals from neutrinos will be much slower. The effective velocity v corresponding to hit i, is then found:

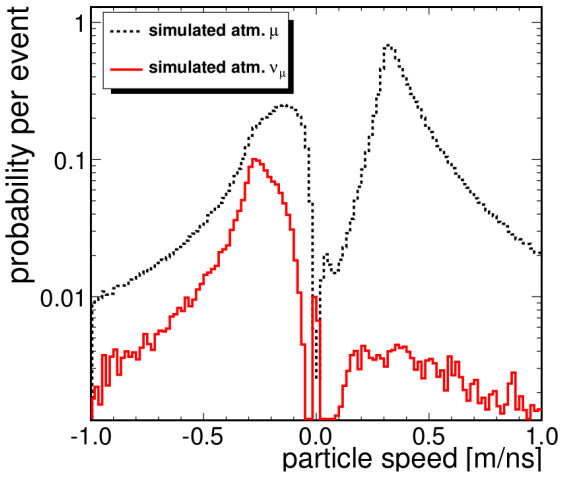


Figure 4.1: Effective particle speed probabilities per event for simulated atmospheric muons (black dashed line) and muon neutrinos (red) inside DeepCore. Figure from [52]. Speeds are positive if the hit occurred before the CoG time and negative vice-versa.

$$v = \frac{||\vec{r}_{\mathrm{CoG}} - \vec{r}_i||}{t_{\mathrm{CoG}} - t_i} \tag{4.3}$$

Atmospheric muons will generally have positive effective velocities because the events begin in the outer regions of IceCube and move inwards, whereas neutrino events will generally have negative effective velocities because the hits start from the central region in DeepCore and move outwards. The distributions are shown in fig. 4.1, and display a clear peak at around $v \approx 0.35$m/ns. Any events containing hits corresponding to effective velocities in the range $v \in [0.25; 0.4]$m/ns are rejected and not included in further processing [52]. Effective velocities above 0.3m/ns are in principle acausal, however, the start time of the event is tied to the center of gravity calculation leading early hits to have overestimated effective velocities. This rejection method retains 99.4% of neutrino events while bringing the atmospheric muon rate down to approximately 17Hz from 280Hz [60].

4.2.3 L2: Processing

The next step in the analysis chain is the standard processing in which low-level variables are calculated. These variables can later be used to define cuts, be added to boosted decision trees, and contribute to further filtering of the event selection. These calculations are part of the standard IceCube software environment and in general they are related to the location, charge and shape of the event. The variables are too numerous to fully list here, instead any variable utilized in later cuts will be described when needed.

4.2.4 L3: DeepCore Atmospheric Muon Rejection

This processing level is part of the standard DeepCore processing, and is mainly concerned with event containment and hits in the veto region. Cuts are implemented using the following variables:

NAbove200

This variable counts the number of uncleaned photo electrons above a height of $z = -200$m and within a time window of 2μs before an SMT3 trigger. For contained neutrino events this is expected to be close to 0 (with possible addition of a small noise contribution). Any event with more than twelve DOMs satisfying this condition is removed.

FirstHitZ

The height coordinate of the first hit in the pulse series can be an indication of whether an event is from an atmospheric muon or a neutrino. Muon events tend to arrive in the downgoing direction and are thus expected to leave hits in the upper part of the detector first. As such, any event with a first HLC hit above $z = -120$m will be vetoed and removed from the sample [60].

RTVeto

The SeededRT algorithm used for hit cleaning can also be used for finding clusters of hits in the veto region in order to identify hits caused by atmospheric muons. In this case the SeededRT is run with a sphere diameter of $r = 250$m and a time window of 1000ns. Any event with a cluster of four or more hits in the veto region is removed [60].

C2QR6

Atmospheric muons generally leave event signatures of long tracks in the detector lasting on the order of 3μs as the muons traverse the detector at close to the speed of light. This yields a relatively constant deposition of charge on the DOMs, which can be used to discriminate the atmospheric muon events from neutrino events that usually deposit their energy over smaller time scales. The *charge fraction* of the total event charge, deposited in the first 600ns is calculated [60]:

$$QR6 = \left. \frac{\sum_{i,t_i} q_{i,t_i}}{Q_{\text{tot}}} \right|_{0<t_i-t_{\text{first}}<600\text{ns}} \tag{4.4}$$

The time is evaluated as related to the first hit in the pulse series, t_{first}. Atmospheric muon events will display $QR6$ values close to zero because they last several times longer than 600ns, while neutrino events will have values close to one. As early noise hits can interfere with this calculation, the first two hits in the pulse series can be ignored to form the *cleaned charge ratio* C2QR6. Any event with C2QR6 < 0.45 is removed from the sample.

4.2.5 L4: Straight cuts + BDT

In similarity with L3, L4 begins with a series of straight cuts, and then trains a *boosted decision tree* (BDT) yielding a BDT score upon which a cut can be placed. The following cuts are put in place:

Time Window Cleaning

A static time window cleaning is positioned around the SMT3 trigger with a range $-3500\text{ns} < t < 4000\text{ns}$, any hits outside this time window are removed. A dynamic time window around the time of maximum pulse density is set with a width of 200ns. Any event with fewer than three hits within this time window is removed [60].

Boosted Decision Tree

A BDT is trained using several variables, some of which were also utilized at L3. FirstHitZ, NAbove200 and QR6/2CQR6 are all included in the L4 BDT. In addition the following variables are used as input:

LineFit Speed

LineFit is a preliminary reconstruction utilized in IceCube. The algorithm fits a plane wave propagating through the ice at velocity $\vec{v}_{\text{LF}}$, to the pulse series. The velocity of said wave can be found analytically [60]:

$$\vec{v}_{\text{LF}} = \frac{\langle t_i \cdot \vec{r}_i \rangle - \langle \vec{r}_i \rangle \langle t_i \rangle}{\langle t_i^2 \rangle - \langle t_i \rangle^2} \tag{4.5}$$

Here $\langle t_i \rangle$ and $\langle \vec{r}_i \rangle$ denotes the average hit time and average hit position respectively. Similarly to other variables, this can discriminate between atmospheric muons and neutrinos, as the more cascade-like neutrino events will tend to have speeds close to zero versus muons that have typical speeds close to $v = 0.3\text{m/ns}$.

Tensor of Inertia

The shape of the event can be quantified via a tensor of inertia calculation akin to the three dimensional tensor of inertia used in classical mechanics, instead of the mass m using the charge q [62]:

$$T_{\text{xx}} = \sum_i q_i (y_i^2 + z_i^2) \tag{4.6}$$

$$T_{\text{yy}} = \sum_i q_i (x_i^2 + z_i^2) \tag{4.7}$$

$$T_{zz} = \sum_i q_i (y_i^2 + x_i^2) \tag{4.8}$$

The relevant variable for distinction of muon and neutrino events is then the eigenvalue ratio:

$$e = \frac{\min(T_j)}{T_{xx} + T_{yy} + T_{zz}}, \tag{4.9}$$

where $\min(T_j)$ represents the smallest moment of inertia along one of the three axes of rotation. Because events caused by atmospheric muons will have most of their moment of inertia centered around one axis of rotation, their eigenvalue ratios are expected to be close to zero. Neutrino events on the other hand are expected to be more symmetric and have an eigenvalue ratio of about 1/3.

BDT Training

A *Boosted Decision Tree* or BDT is a learning algorithm to classify events based on multiple variables. The BDT is trained using MC simulation for which the true values are known, and returns an overall assessment of the 'signalness' of an event based on the input variables in the form of a score ranging from -1 (background like) to +1 (signal like). While straight cuts can work well in some cases, generally speaking not all background events will be very background-like in all variables, while not all signal events will be very signal-like in all variables. Therefore, a BDT can give an event a high signal-like score even if some variables are indicative of background, but other variables are indicative of signal. This score is calculated for each event and then used to enforce a standard straight cut. For the GRECO selection all of the above variables are included in the BDT and events with BDT scores below 0.04 are excluded [60].

4.2.6 L5: BDT Background Removal

Another BDT is trained at this level, using a set of variables consisting of: T75Q, VICH, Radial Distance, Quartiles_CoG, Z-travel, and SPE-zenith. A full description can be found in [60], but the variables are briefly discussed in the following.

Quartiles Center of Gravity Distance

The pulse series is divided into its four quartiles in time. The center of gravity is then calculated for both the first and the last quartile. The distance between them is then the variable used. It is a meaningful variable as low energy neutrino events will have short distances between their first and last centers of gravity, while atmospheric muons typically traverse longer distances in the detector and will show longer distances between the first and last CoG.

Radial Distance

Like the position of the first hit in the z-direction was used to veto downgoing atmospheric muon events, so can the distance from the center of IceCube at string 36 be used to veto signals starting in the outer layers. This is measured by the *radial distance* variable ρ of the first hit in the pulse series:

$$\rho_{36} = \sqrt{(x - x_{36})^2 + (y - y_{36})^2} \tag{4.10}$$

The radial distance is utilized in the training of the level 5 BDT.

SPE Zenith

The goal of this *Single Photo electron* reconstruction is to identify downgoing muons by reconstructing the direction of the event. The algorithm assumes each DOM is hit by just one photon, and performs a likelihood fit. These are usually time consuming, however, at this point in the selection chain the event rate is low enough that running such a fit is plausible without severely impacting computation time [60]. The time for a Cherenkov photon to arrive at an IceCube DOM from a single emission point can be calculated as:

$$t_{\text{point}} = t_{\text{emit}} + \frac{n}{c}|\vec{r}| \tag{4.11}$$

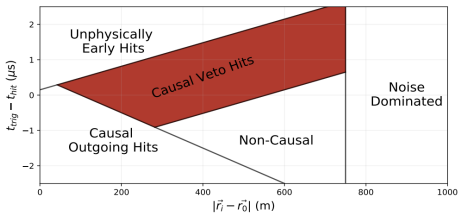


Figure 4.2: Schematic of the causal regions connected to the Veto ID Causal Hits algorithm.

with $\vec{r}$ being the vector from the point of emission to the DOM, n the spectral index of the ice, c the speed of light in vacuum and t_{emit} the photon emission time. Likewise, the corresponding calculation can be done for Cherenkov photons from a muon track [65]:

$$t_{\text{track}} = t_{\text{emit}} + \frac{\vec{r} \cdot \hat{n} + \rho \tan(\theta_C)}{c} \tag{4.12}$$

where $\hat{n}$ is the unit vector pointing along the track direction, ρ is the impact parameter of the track as incident on the DOM: $\rho = |\vec{r} - (\vec{r} \cdot \hat{n})\hat{n}|$. Lastly θ_C is the opening angle of the Cherenkov cone, also known as the Cherenkov angle. These formulas are idealized and ignore any contribution to the observation time due to scattering of photons in the propagation medium. Any such scattering will lengthen the travel path distance of the photons, delay their arrival time as compared to the expectation, and give rise to a *time residual distribution.* Unfortunately no analytical form has been found, but previous work in AMANDA [62] , found an approximation in the *Pandel function:* $P(\Delta t_i | \vec{r}_{\text{vertex}}, t_{\text{vertex}}, \hat{n})$, with $\Delta t_i = t_i - t_{\text{point}}$ [66], which estimated the scattering as a function of distance between the DOM and the emission point. A likelihood can then be constructed:

$$L(\vec{r}_{\text{vertex}}, t_{\text{vertex}}, \hat{n}) = \prod_{i}^{\text{pulses}} \frac{dP_{\text{Pandel}}(\Delta t_i | \vec{r}_{\text{vertex}}, t_{\text{vertex}}, \hat{n})}{dt} \tag{4.13}$$

This likelihood can then be maximized to obtain a best-fit value for the spacetime position and direction of the track. This reconstruction relies on the assumption of an infinitely long muon track without starting and stopping points. This is a reasonable assumption for a muon passing through IceCube. The resulting reconstructed zenith angle is used in the L5 BDT, as atmospheric muons are predominantly downgoing.

T75Q

The *T75Q* variable represents the time to reach 75% of the total charge in an event. At first glance this is related to the C2QR6 variable. However, where the former searches for a charge within a time window, the T75Q does the opposite and searches for a time as a function of charge. The physical mechanism between the discriminative power is also similar: atmospheric muons travel longer and will take longer to deposit 75% of the charge in the event as compared to low-energy ($E < 100$GeV) neutrinos which will deposit their energy quickly

VICH

The *Veto ID Causal Hits* algorithm finds hits in causal contact with the first DOM hit that led to the SMT3 trigger [63]. It works on an uncleaned hit series and removes hits based on sorting into the different causal regions shown in fig. 4.2. These include hits outside causal contact with the first trigger hit, either due to occurring before the trigger or to lying outside the causal light cone. Also excluded are noise hits, and hits too far from DeepCore to meaningfully contribute to the event even though they are in principle in causal contact. The algorithm then returns the total charge collected from all the remaining hits, which are marked in red in fig. 4.2.

Z Travel

Atmospheric muons do not carry enough energy to penetrate far through the Earth. Typical distances are on the scale of a few kilometers, which is why all atmospheric muon events are expected to be in the downgoing direction. In order to estimate the direction of travel without doing a full reconstruction of the event the *Z Travel* calculation is implemented. Like the quartiles distance variable, the hits included in the first time quartile are used to calculate the center of gravity for that set of hits. This CoG is then used as a starting point, and the Z Travel variable is calculated as the charge weighted average distance of hits in the z direction only.

$$\Delta z = \frac{\sum_i q_i (z_i - z_{\mathrm{CoG}})}{\sum_i q_i} \tag{4.14}$$

Seeing as the neutrino signal is expected to be approximately isotropic save for oscillation/matter/horizon effects, the associated Δz values are expected to fall in both the positive and negative regime, while being relatively small compared to the detector size due to their typically low energy. On the other hand, atmospheric muons are expected to be downgoing, so their Δz values are expected to be negative and large as compared to the typical values for the neutrinos.

L5 BDT

Using the variables described above, another BDT is trained. After each training step the least important variable is removed in an attempt to curtail overtraining. The process is repeated until only six variables remain [60]. A cut is placed at the BDT value of 0.04 yielding a 95% reduction in atmospheric muons for only a 30% reduction in the neutrino rate. While this is a significant reduction, it does have the benefit of lowering total computation time for the complex calculation performed at Levels 6 and 7.

4.2.7 L6: Accidental Event Removal and Corridor Cut

Corridor Cut

Even with the above steps in the selection chain to limit the number of atmospheric muons, a few are still present in the sample [60]. These are the elusive atmospheric muons that enter IceCube along just the right trajectory to avoid detection until they reach the densely instrumented inner region. This can happen due to the geometry of the strings and the placement of the DOMs. The hexagonal pattern allows for straight 'corridors' of low detection sensitivity along which atmospheric muons can propagate undetected, as shown in figure 4.3. The corridor cut aims to diminish this effect by looking for SLC hits along these corridors that are correlated with hits in DeepCore [59]. The center of gravity

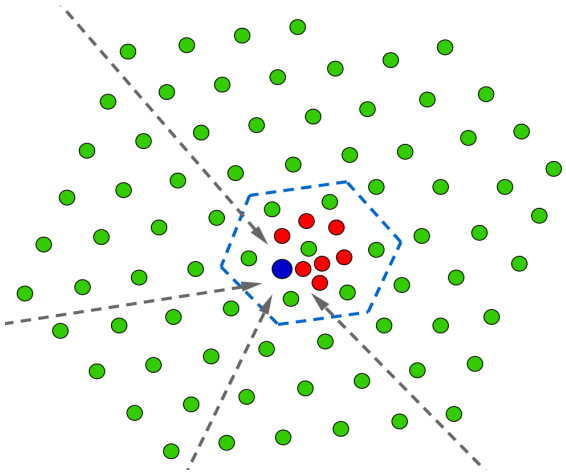


Figure 4.3: 'Blind corridors' allow atmospheric muons to reach the inner detector region undetected. The blue dot marks the first hit in this example event, with red dots marking other hit strings in the event. Arrows indicate the 'blind' directions for the event in question. Figure from [59].

is calculated for the event and the hit string closest to the CoG is used to determine which strings to check for 'corridor' hits. The algorithm then returns the number of hits found. While in the idealized case a value of zero would be favorable, in reality the occurrence of noise hits will often cause correlated corridor hits, so a higher value is necessary to avoid severe signal suppression. The scheme thus allows one hit on the corridor strings, while events with two or three hits in said region will be removed.

Fill Ratio

The *Fill Ratio* algorithm was developed as a search for high energy cascades and relies on a reconstructed vertex and a pulse series as input. It calculates the average distance from the interaction vertex to all the hits included in the pulse series [67, 60]. It then finds the fraction of DOMs hit within this average distance 'containment' sphere, thereby quantifying the density of the event. Track-like events are expected to show low fill ratios because they are elongated and a substantial part of the DOMs in the containment sphere will not have registered any hits. For cascade events the light distribution is expected to be roughly isotropic, so the FillRatio is expected to be close to one. These assumptions have been tested and the algorithm implemented for the high energy IceCube neutrino searches in the energy region above O(10) TeV, where it has proven to be an effective discriminator between track and cascade events [67]. However, for the O(10) GeV events of interest in DeepCore this discriminating power all but vanishes due to the short muon tracks exhibited at these energies. Nonetheless, the FillRatio has been shown to have significant power to suppress accidental events arising from detector noise [60]. Accidental noise triggers are generally characterized by a roughly uniform distribution of hits in the detector, as opposed to neutrino events, which tend to have hits in clusters around the interaction vertex. As such the FillRatio calculation for these events will have a large average radius and a small value of the FillRatio variable. Setting the FillRatio cut at a value of 0.05 and removing events with FillRatios lower than this yields an order of magnitude reduction in the accidental background, while only reducing the neutrino signal by about 10% [60]. Importantly,

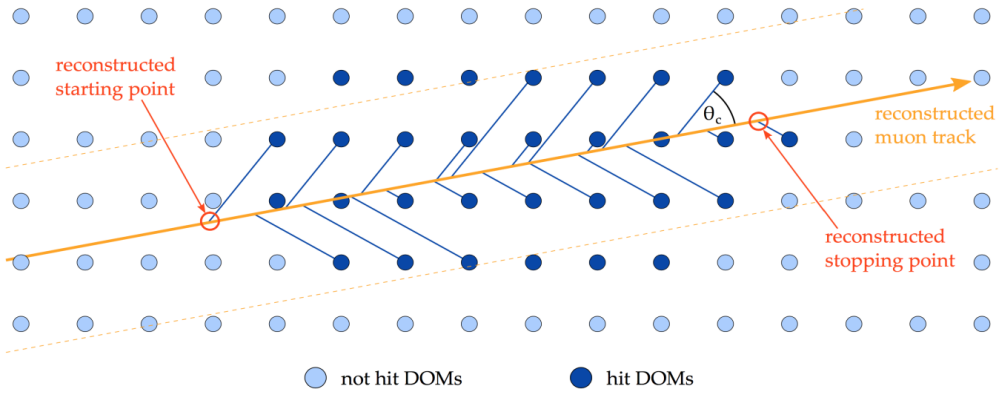


Figure 4.4: FiniteReco reconstruction of the start- and stopping points, from a reconstructed muon track direction. Figure from [63].

with this cut on the FillRatio the over all accidental noise trigger rate drops below the neutrino trigger rate.

FiniteReco Containment

The *FiniteReco* reconstruction algorithm is a refinement of the infinite track reconstruction used in the SPE-Fit reconstruction discussed above [63]. Where SPE-Fit assumes an infinite muon track, FiniteReco aims to set meaningful limits for said track. This is done by starting from a seed consisting of the previous SPE-Fit reconstruction and then assuming light emission from the muon propagation path along the Cherenkov angle, to find the start- and end points of the track. This is illustrated in fig 4.4. FiniteReco does not alter the direction of the track from the seed, but estimates the location of the interaction vertex from the reconstructed starting point [63]. If an event is outside the DeepCore fiducial volume it is removed from the sample [60].

4.2.8 L7: Event Reconstruction

The basic data measured by IceCube is charge- and timing information in the form of pulses registered on the DOMs. The data gathered from the detector can be used to infer information about the physical properties of the interaction in the ice. Specifically, the energy and direction of the incoming particle are of interest. The reconstruction algorithm used in the GRECO sample is known as Pegleg and was developed in IceCube by the PhD student Martin Leuerman [58]. Pegleg makes use of two other reconstruction algorithms, Monopod and Lilliput (stable IceCube software) and will always attempt reconstruction of a hybrid event, consisting both of a track and a cascade. The track is reconstructed in segments of five meters, while the cascade is reconstructed based on a center of gravity for the DOM hits not associated with the track. Pegleg fits eight parameters: The position (x, y, z), time (t), direction in zenith and azimuth angles (θ, ϕ), the track length (L), and the energy in the cascade (E_{cascade}). The total energy of the event is calculated by assuming the track part comes from a minimally ionizing muon with an energy loss of 220MeV/m in the ice:

$$E_{\text{total}} = E_{\text{cascade}} + L \cdot 220\text{MeV/m} \tag{4.15}$$

Pegleg starts from an event hypothesis at an arbitrary point in this parameter space, and can be seeded with the results from a preliminary fit using Monopod. The algorithm works on a collection of DOM hits, including time and charge information. From the event hypothesis the light output is estimated using photonic spline tables based on the ice properties. Pegleg then calculates a likelihood for each DOM to observe a given charge and finds the total likelihood as the product of all individual likelihoods for all DOMs. This likelihood space is then organized into time slices and minimized independently for each time step using the Multinest minimizer package [59]. The output is used to estimate the underlying true likelihood space in order to create new event hypotheses, which are then in turn minimized.

Fitting in this eight dimensional parameter space requires substantial computing power for processing. In order to bring down the computational requirements, simplifications are introduced. A discretization of the track length is set to the track length step size used when generating splines based on the ice properties. Only DOMs within 150m of the particle hypothesis position are expected to carry significant charge. All other DOMs are considered to have charge levels compatible with background noise. This speeds up the algorithm by avoiding computationally expensive estimates of charge on distant DOMs, but comes at the price of less energy resolution.

At L7 there are extra optional containment cuts, which are not included for the sample used in this work. These were introduced in order to minimize the disagreement between MC and data, and came at a loss in the neutrino event rate [60]. While such disagreement is critical in the case of forward fitting analyses, the unfolding method used in this dissertation relies more heavily on statistical knowledge of the response matrix. For each simulated event the MC truth is correlated with the reconstructed value to give an entry in the response matrix. This type of unfolding analysis is more sensitive to uncertainty in the response matrix than it is to background contamination. This will be further investigated in section 6.6. While disagreement between data and MC is never preferable, the iterative unfolding method can compensate for the effect because it converges the unfolded result on the most likely true value with the optimal number of iterations. The method is described in detail in chapter 5.

The detector resolution in energy and direction defines two dimensions in the response matrix, with a third dimension being particle identification. The smearing in energy and direction due to detector resolution, stemming from the precision of the reconstruction algorithm, will be mitigated by the unfolding. The unfolding machinery takes as input the expected signal, the preparation of which is the topic of the rest of this chapter.

4.3 Detection Efficiency

The detection efficiency describes the fraction of interacting events detected by IceCube. This efficiency is defined as the probability for a neutrino interacting in or near the instrumented ice volume to be included in the data sample at final level. The efficiency is measured via an MC study as described in the following. The base neutrino simulation data, known as generator level data (or L0 data), is compared to the GRECO final level sample (Level 7, L7), and the efficiency is defined as the ratio of events N between the two:

$$\epsilon = \frac{N_{L7}}{N_{L0}} \tag{4.16}$$

This definition of the efficiency can thus be used to translate back from a measured number of neutrinos to the number of neutrinos interacting in the ice. From a technical standpoint the L0 data is re-weighted in a standard way to reflect the relative strengths by flavour of atmospheric oscillations, cross sections, and flux while removing the fiducial volume and livetime. By doing this, the L0 data sample is one-to-one comparable with our standard simulation after all selection cuts (L7). Normally an efficiency calculation presents a binomial problem, but due to the re-weighting this is not entirely true in this case and special care must be taken when calculating the statistical uncertainty.

4.3.1 Statistical Uncertainty on the Efficiency

Since the efficiency is calculated from weighted distributions it would be wrong to assume the standard binomial relationship between L0 and L7 data. Both are weighted the same way, so in order to determine the uncertainty we must make use of the uncertainty for a weighted binomial distribution:

$$\sigma_\epsilon = \frac{\left[(1-2w)\,\sigma_{L7}^2 + w^2\sigma_{L0}^2\right]}{N_{L0}^2} \tag{4.17}$$

, with $w = \frac{N_{L7}}{N_{L0}}$,σ_{L7} and σ_{L0} being the uncertainty on the L7 and L0 samples respectively. The calculation is confirmed to be consistent between manual calculation as well as the method implemented in ROOT.

4.3.2 Overall Detection Efficiencies

The detection efficiencies for each unfolded channel have been determined using the method described above, with the results presented in fig. 4.5. Generally the efficiency is on the few percent level, with some bins reaching eight or nine percent. This is generally consistent with the expectation based on the impact on the event rate from the different data sample processing levels [60].

4.4 Background: Atmospheric Muons

Muons generated from cosmic ray interactions in the atmosphere constitute the dominant background in this analysis. As discussed in sec. 2.6, muons can travel tens of kilometers before having lost all their energy, and will leave clear Cherenkov tracks in IceCube/DeepCore. Several of the cuts described in section 4.2 are designed for the sole purpose of bringing down the rate of atmospheric muons included at final sample selection level. The simulation of atmospheric muons is performed using both Corsika [64] and MuonGun [60]. While the Corsika muon simulations are produced for the IceCube collaboration as a whole, the selection cuts in sec. 4.2 reduce the standard simulation Corsika background to such an extent that it is statistically meaningless, and a new simulation scheme dubbed *MuonGun* was developed for use in the low-energy regime below 1 TeV [60]. MuonGun generates muons on a cylinder surface and intentionally aims them at a specific region of IceCube. This introduces some bias to the simulation with certain regions being slightly over simulated and others being under simulated. Particularly, it can be configured to increase statistics in the region of interest, namely DeepCore, while accepting lower statistics in the outer regions of IceCube. The *generation cylinder* measures 1600m in height with a radius of 800m, thereby optimizing the volume to surface ratio. The spectrum from which the muons are generated is customizable, and is drawn from an 'offset' energy power law, for historical reasons [60, 65].

$$f(E) = (E + E_0)^\gamma \tag{4.18}$$

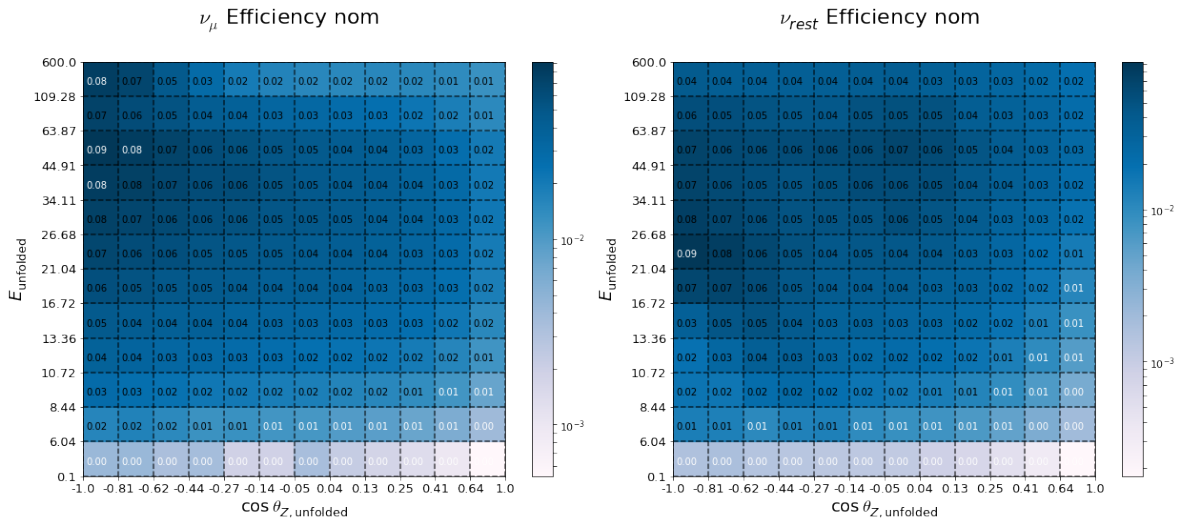


Figure 4.5: Efficiency of the combined trigger- and selection efficiency for the $\nu_\mu^{cc} + \bar{\nu}_\mu^{cc}$ channel (left) and the ν_{rest} channel (right) respectively.

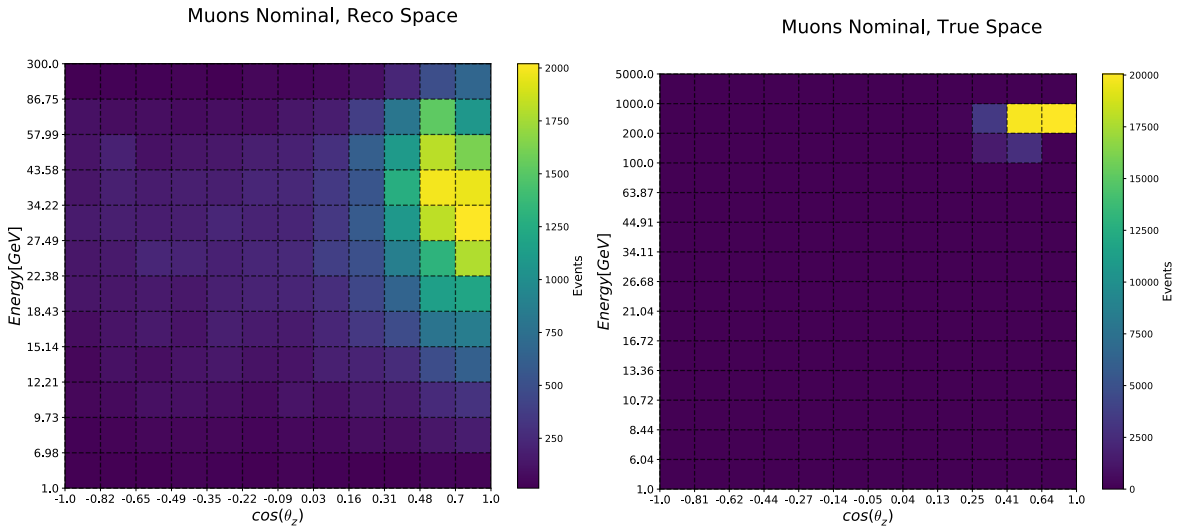


Figure 4.6: Atmospheric muons in the final sample. Left: Reconstructed space. Right: True space. Note how the reconstruction smears the otherwise well confined muon signal in true space into a large part of the reconstructed parameter space. Note the color scale is a factor of ten times higher in the right plot.

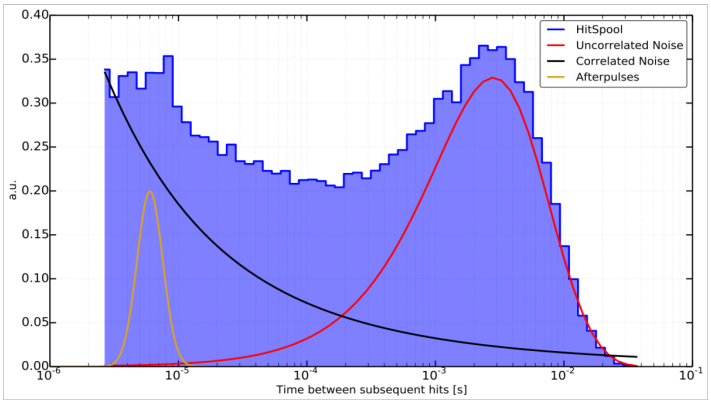


Figure 4.7: HitSpool data for DOM 15 on string 27, shown in blue. The various contributions to the HitSpool signal are shown. Red: Poissonian noise (un-correlated). Black: non-poissonian noise (correlated). Yellow: Afterpulses. Components are not to scale. Figure from [60]

E is the energy of the muon, E_0 an energy offset and γ the spectral index. For the simulation used for the GRECO sample simulation, the offset energy was set to $E_0 = 700\text{GeV}$ with a spectral index of $\gamma = -5$. This is a notable difference to the spectral index of the cosmic rays at about $\gamma = 2.7$ [6]. The steep spectrum is selected in order to facilitate generation of lower energy events, greatly reducing the computation time. Furthermore, low energy muons produce little light in DeepCore and are difficult to identify. It is therefore expected that low energy muons will be the dominant component of the background at final sample selection level. While the atmospheric muon flux is correlated with the atmospheric neutrino flux to some extent, the MuonGun simulation breaks this degeneracy and allows the muon spectrum to float. The simulated muon spectrum at final sample selection level including both Corsika and MuonGun components is shown in fig. 4.6, and indicates that the muon background contribution is concentrated in the downgoing region at energies between $\sim 20\text{GeV} - 80\text{GeV}$. In true space the muon contribution is strictly downgoing and at energies above $\sim 100\text{GeV}$. The overall muon background contamination is about 26%, and while this may seem large, the unfolding will to some extent be able to mitigate the effects of the background, as will be discussed in chapter 5. First, another background needs to be discussed: The detector noise.

4.5 Background: Noise

Random noise events in the DOMs arise from thermal radiation and radioactive impurities in the metals contained in the DOM electronics, PMT glass, and DOM pressure vessel. This has been measured by extracting the '*HitSpool*' data (data taken directly from DOMs without a trigger). Once a radioactive decay happens, a rapid succession of hits are recorded on the PMT, with the signal lasting up to a few milliseconds, as shown in fig. 4.7 [68]. The hits are thought to be caused either by scintillation or luminescence processes. The overall noise rate is about 540Hz for IceCube DOMs and 780Hz for the high-QE DeepCore DOMs.

The *thermal noise* is temperature dependent and stabilizes over the seasons following a slow exponential decay, and typically occurs with a frequency around 200Hz. The *radioactive decay noise* on

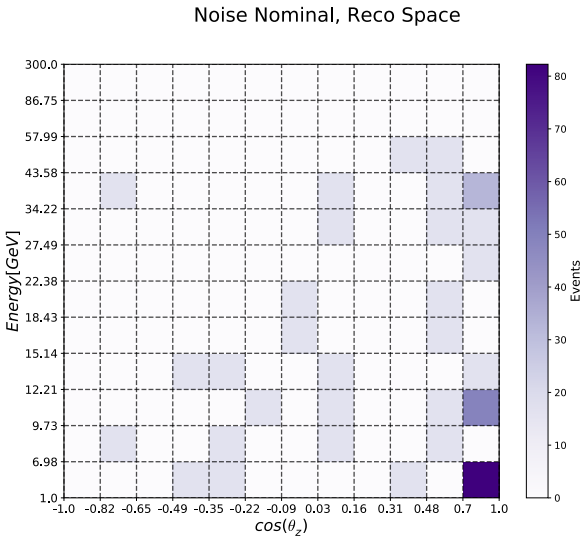


Figure 4.8: Noise contribution at final level. Notice that noise is not considered to have a smearing effect, but simply contribute to the event rates - hence the distributions will be identical in Energy and $\cos(\theta_z)$ for both reconstructed space and true space. The will however, in the true space, be sorted into the background bin.

the other hand typically has a value around 50-100 Hz. Both are simulated by the IceCube software module known as *Vuvuzela*, and fits both a Poisson and non-poisson component to the hit distribution individually from each DOM. The noise model and the Vuvuzela software module was developed by M.Larson as his Master's Thesis [68]. Applying the noise simulation to the current data sample with the correct livetime results in the event expectation shown in fig. 4.8. The energy and $\cos(\theta_z)$ here refer to the reconstructed energy and reconstructed event direction respectively, as noise events do not contain true particle information in the MC for these variables. For the unfolding analysis presented in this work, the background signal is classified as such and sorted into the background bin on the truth side of the unfolding matrix.

5 Analysis Method

Unfolding is the process of removing systematic detector effects from a given measurement, in order to estimate the true distribution giving rise to the measurement. It is particularly useful in the cases where underlying shifts between true values of nature and measured values are observed. It is of course not immediately clear when this occurs, due to the nature of the problem. This can for example occur when the systematic shift is not constant but depends on other variables, leading the shift to follow a non-trivial distribution. In general this effect is known as instrumental smearing. The challenge then, in an unfolding analysis, is to estimate this smearing and compensate for it. In the following we will discuss the unfolding method used.

5.1 Unfolding vs Forward Folding

Two main types of 'folding' analyses are widely used in high energy physics; Forward folding and unfolding. Here we will briefly touch upon the merits of each, and finally do a qualitative comparison. We assume a measurement is represented by a dataset D and strive to find information about the underlying true distribution, under the assumption that our measurement instruments are inherently inaccurate.

In the forward folding approach this is done by assuming a model, and from that a Monte Carlo simulation set is built, which then represents the model truth. This estimated truth is then folded with the smearing of our instruments, also known as the detector response. It is then possible do a one-to-one comparison with the observed data. This process is then repeated for many different models and the one which fits best to the data is then presented as the result. The forward folding approach thus gives us the model which best explains the observed data and is commonly used for parameter determination. The process is illustrated as going from left to right in fig. 5.1.

The unfolding approach reverses this process: starting from the observed data, the detector response is removed from the measured data set to yield an estimate of the underlying true spectrum. Except for the knowledge of the detector response, the unfolding approach makes no assumption regarding the nature or structure of the data or the underlying distribution. The unfolding method thus gives us the best estimate of the underlying truth in a model independent way, and is illustrated progressing from right to left in fig. 5.1.

Both methods rely on extensive knowledge of the detector response, usually obtained via Monte Carlo simulations, and both are biased to some degree. The forward folding method focuses on specific models, with parameters which can be determined to a certain level of precision. However, measurements are always uncertain, so many different models with different parameters could possibly describe the data. Occam's razor prescribes that we focus on the simplest models, so usually forward folding analyses test only on one specific model. As such they tell us nothing about the merits of other models, or whether the assumed truth to which the smearing is applied is actually the one realized in nature. In the context of this analysis a forward folding method would measure a reconstructed event rate and then use a model to fit, with flux as a parameter.

On the other hand, the unfolding approach attempts to be model independent by removing the instrumental effects from the measurement. As knowledge of the detector response depends on detector calibration and MC simulation, there is still a risk of being model-biased. Getting rid of this bias is one of the challenges when doing an unfolding analysis. Provided this bias is overcome the unfolding gives an accurate estimate of the underlying truth from which the data sample was generated. This, however, gives no indication as to the model description of the underlying truth. In the context of this analysis, in

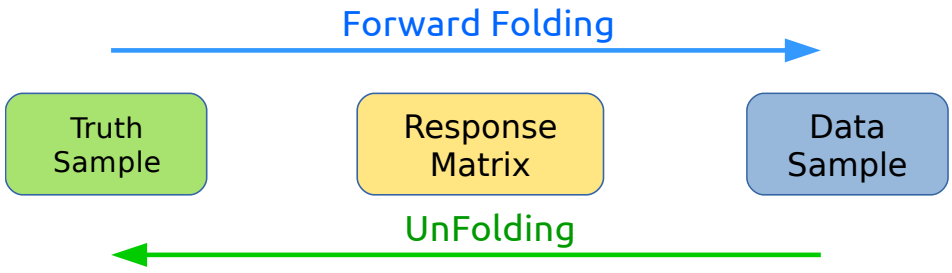


Figure 5.1: Forward folding versus unfolding in a schematic form. Forward folding techniques start from a truth sample in the form of simulated data, smear that via the response matrix and compare with the measured data. Unfolding reverses the process: Starting from the data sample, removing the effects of the response matrix and end up with a truth sample, estimating the underlying truth of nature.

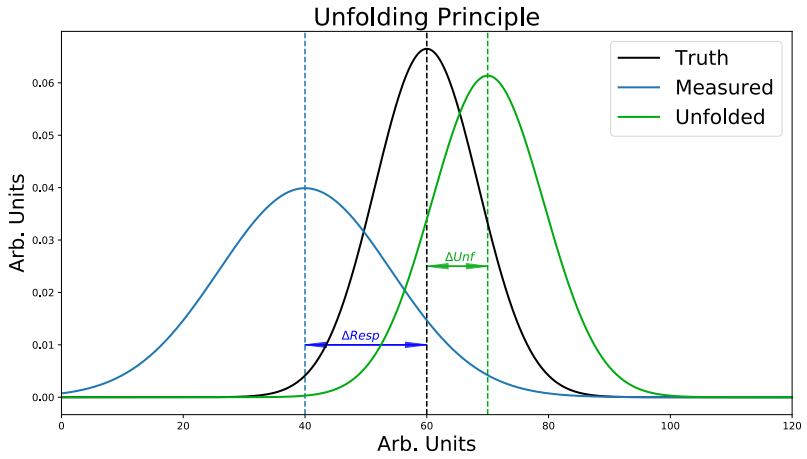


Figure 5.2: Illustration of the unfolding conceptual procedure. Black represents the population and blue the sample. They are related by the response matrix, which gives rise to the shift between the population and sample means in Δresp. The unfolded histogram is an imperfect reversal of that process and is shown in the green curve. The associated error (sometimes known as the unfolding bias) is illustrated by Δunf.

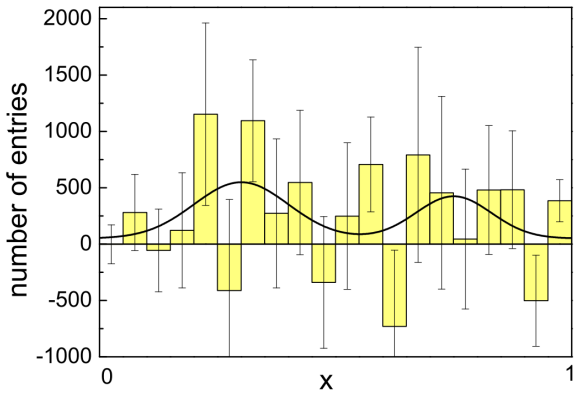


Figure 5.3: Example of 'oscillatory' behaviour in a Matrix Inversion unfolding, without sufficient regularization. [69]

order to quantify how many neutrinos pass through IceCube we can measure a reconstructed event rate, apply the unfolding method, and find the underlying true event rate by unit volume ice without fitting for a parameter. This unfolding analysis will proceed in multiple dimensions, making simple graphical illustrations a challenge. Instead, we will employ illustrative schematics in order to introduce concepts in a comprehensible way throughout. An example of an idealized unfolding procedure is illustrated in fig. 5.2. A true distribution (commonly known as the population in statistics) is shown in black. Through reconstruction uncertainties and detector inefficiencies, at measurement, the average value and the width of the distribution may be altered, and is shown in the blue curve (commonly known as the sample in statistics). The mean value shift ΔResp and widening are described by the response matrix, which maps true values onto reconstructed values. The main task of the unfolding process, is then to recover the population mean and variance from the sample distribution. However, not all unfolding methods are perfect and an error may be introduced, prompting the unfolded distribution to be shifted away from the true distribution by ΔUnf. In MC closure tests it is possible to quantify this error, however, it is impossible in a real data unfolding. It is therefore necessary to explore the parameter space of various combinations of input and response matrix to test the consistency of the unfolding. Several tests have been developed and will be presented in chapter 8.

Forward folding and unfolding methods each come with their own advantages and disadvantages. Forward folding is useful for specific model testing or differentiating between various models. Unfolding on the other hand is ideal for model independent measurements of physical quantities. One advantage of unfolded results is they can be used as inputs in other contexts without the knowledge of the detector response. Many different unfolding techniques exist, and in the following we will discuss the choice of unfolding method used in this analysis.

5.2 Unfolding Method

Several methods of unfolding exist and among the candidates are such implementations as Single Value Decomposition (SVD) , Matrix Inversion, Iterative Bayesian Unfolding or bin-by-bin unfolding [11]. A short motivation for our choice of method will be given here. In general a good unfolding method should live up to the following criteria:

1. Does not introduce a bias on the unfolded result from the input MC
2. Uses as much information as possible (for example works in multiple dimensions)
3. Is independent of prior expectations (model agnostic with respect to eg. smoothness or shape)
4. Accurately describes the true distribution in closure tests

While many methods have similar advantages each comes with its own drawbacks: Bin-by-bin is significantly biased towards the MC [69]. Both SVD and Matrix Inversion will yield some fluctuations in the unfolded spectrum (known as "oscillatory behaviour"), which must be corrected for via regularization, an example of which is shown in fig. 5.3 [69]. Iterative Bayesian unfolding risks a bias towards the MC in the case of few iterations, but can yield large uncertainties at high iterations [70].

In this analysis the the D'Agostini iterative Bayesian unfolding method was chosen, and used the version implemented in the RooUnfold software package for ROOT [72]. The software can handle multiple dimensions and the non-trivial task of implementing an expansive regularization is avoided. While the initial iterations are biased towards the MC this bias can be overcome with sufficient iterations. The method is a maximum likelihood estimator, and calculates an unfolding matrix starting off from Bayes theorem:

$$P_n(i|j) = \frac{P(j|i) \times P_n(i)}{P_n(j)} \tag{5.1}$$

Here $P(i|j)$ on the left hand side is the unfolding matrix, giving the probability of a measurement in bin j arising from a true signal in bin i. The other terms can be described as follows:

- $P(j|i)$ is the response matrix and gives the probability for a true event in bin i to yield a measurement in reconstructed bin j. It represents our knowledge of the detector hardware response and the reconstruction algorithms. It describes the impact of the detector smearing on the measurement and can generally be estimated from detailed Monte Carlo simulations.
- $P_n(i)$ is the probability of a true event to occur in bin i. It thus represents the prior knowledge of our true signal. For the initial value of this, we use a model spectrum and then calculate:

  $$P_0(i) = \frac{N_i^{mc}}{\sum_k N_k^{mc}} \tag{5.2}$$

  As this term is built from an a-priori model it can introduce a bias in our analysis, however, optimizing the number of unfolding iterations will compensate for this effect.
- $P_n(j)$ is the probability for an event to occur in reconstructed bin j, and is typically called the marginalized probability in Bayesian analysis. It contains the full information regarding the parameter space, which then corresponds to the sum over all probabilities of the true bins i contributing to that reconstruction bin j:

  $$P_n(j) = \sum_i P(j|i) \times P_n(i) \tag{5.3}$$

Using the Bayesian approach we can then make an estimate of the true energy spectrum U_i^n simply by applying the unfolding matrix to the relevant bin in our measured spectrum m_j:

$$U_i^n = P_n(i|j) \times m_j \tag{5.4}$$

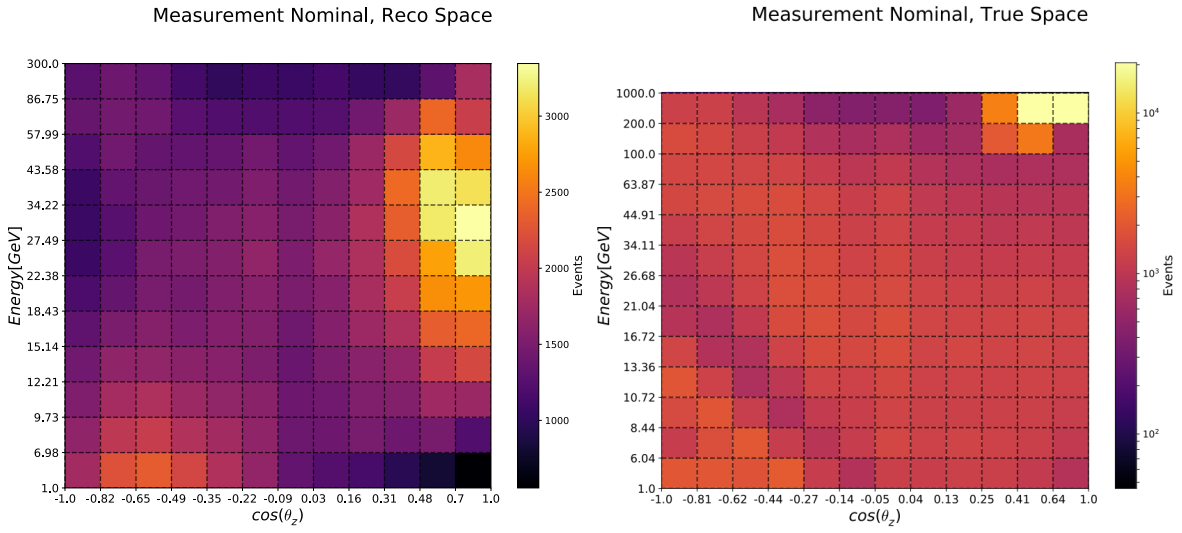


Figure 5.4: Events at final processing level. Left: Reconstructed Space. Right: True Space.

U_i^1 then serves as a first estimate of the true distribution. At this point the prior we introduced can have a strong influence on the unfolded distribution, but we can move the unfolding away from the prior by iterating the whole process. Ie: we take the unfolded spectrum as the input prior for the next iteration and generate new prior and marginalized probability terms. We then calculate a new unfolding matrix $P_2(i|j)$ and apply it to our measurement to get a less biased estimate of the truth U_i^2 . The iterative process works under the assumption that applying the unfolding matrix to our data brings the unfolded distribution closer to the truth - in essence we move the prior term closer and closer to the truth. This is unfortunately not always the case, as statistical fluctuations will gain increasing weight with the number of iterations and may eventually pull the unfolded distribution away from the truth as the number of iterations increases. In order to combat this 'overfitting', a reasonable number of iterations must be estimated, which is explained in the next section.

5.3 Expected Signal

The total number of simulated events that make it through to final level is 491094. These events are weighted to fit the expectation from the atmospheric flux model and based on the genie cross sections, giving 170461.3 expected signal events. In addition there is a background contribution of 43400.1 muon and noise events, making for a total expectation value of 213861.4 events. With this the background contamination is expected to be 20.1%.

IceCube/DeepCore can identify muon neutrino CC events, but ν_e^{cc} and ν_τ^{cc} events will often be indistinguishable (except in the case of the "double bang" event signature), and as such will be grouped together. NC events are indistinguishable amongst flavors by nature, and are strikingly similar to the ν_e^{cc} and ν_τ^{cc} events, thus these event types will be grouped together. The expected neutrino rates are summarized in tab. 5.1.

Event Type	Events	Rate [mHz]
Neutrino	170461.3	1.551
Background	43400.1	0.395
Total	213861.4	1.946

Event Type	$\nu_\mu^{cc} + \bar{\nu}_\mu^{cc}$	ν_{rest}	Atm. μ	Noise
Expectation	110040.8	60420.5	42852.5	547.6

Table 5.1: Expected number of events and event rates for neutrinos and background. Notice these are expectation values, and can thus be non-integer.

5.4 Iterations

As we increase the number of iterations in the unfolding, we put less emphasis on our prior, but with many iterations eventually run the risk of putting too much emphasis on statistical fluctuations. The number of iterations used in an analysis can then seem like an arbitrary choice, but here we will describe methods to find the optimum number of iterations. When building the unfolding matrix, we make extensive use of our Monte Carlo simulation. The simulated events are generated individually from a power law, re-weighted to match the Honda flux model and IceCube effective area, and run through the Pegleg reconstruction. As such we have full information regarding both the MC truth and MC reconstructed spectra. From the weighted MC, we generate a pseudo experiment as a Poisson variation around the nominal MC (the MC generated with all systematics set to their nominal values). In these tests, this pseudo experiment plays the role of our measured data, while the weighted MC input to the response matrix, prior and normalization terms are used they way they will be after unblinding of the measured data. The pseudo experiment is thus statistically independent from the weighted MC. For simplicity, here the livetime, volume and efficiency are not compensated for and the unfolding test proceeds from reconstructed events to unfolded events. We apply the D'Agostini iterative unfolding approach to the pseudo experiment, and since we know the exact truth spectrum from which the pseudo experiment was generated, we can compare the two via a simple $\chi 2$ test:

$$\chi^2 = \sum_i \left(\frac{U_i - T_i}{\sigma} \right)^2, \tag{5.5}$$

where U_i represents the bins in the unfolded pseudo experiment spectrum, T_i the corresponding bins in the MC truth spectrum, and σ is the statistical uncertainty on the bin in question (calculated via propagating the standard statistical uncertainty through the unfolding). We can now let the unfolding algorithm iterate, calculate this chi square for each step and stop the iterations once the minimum chi square value has been found. The optimal number of iterations will strongly depend on the effect of the systematic uncertainties. For simplicity here we include a simple example in which we test three different scenarios:

1. Asimov test: a pseudo-Asimov test in which we do not use a pseudo experiment as input, but instead use the values from the reconstructed MC directly. Results for the test statistic are shown in fig. 5.6

2. Pseudo experiment test: a pseudo experiment as described above. The test statistic is averaged over 200 trials each unfolded separately. The toy model spectra and unfolded results are shown for 1 and 30 iterations in fig. 5.5, while the evolution of the test statistic is shown in fig. 5.7

3. Systematic uncertainty test: 200 pseudo experiments as above but with the energy spectrum shifted by a power law index of +/- 0.03, to emulate the impact of such a systematic uncertainty. The unfolded results are shown in fig. 5.8 with the evolution of the test statistic shown in fig. 5.9

In the pseudo Asimov and the pseudo experiment cases we get good agreement already after the first iteration, since the pseudo measurement is very close to what the unfolding algorithm is built from. This then constitutes a successful closure test of number of iterations. In the case of the systematic uncertainty test with a spectral shift we see the following: as expected, a rapid decrease in the test statistic with number of iterations is observed, after which a plateau is reached, before the statistical variations start getting too much weight and again pull the chi square to higher values. This is results

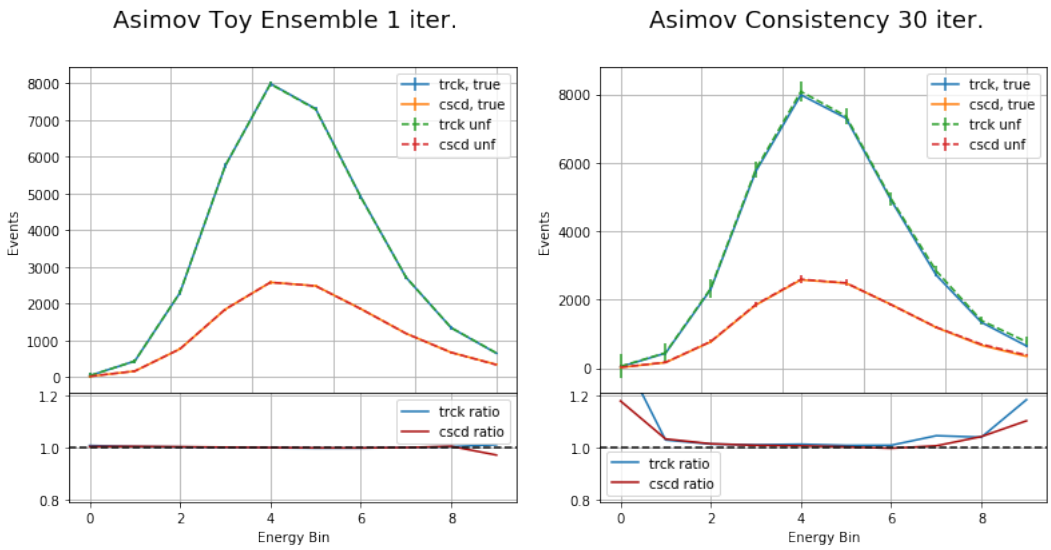


Figure 5.5: True and Unfolded spectra for the Asimov unfolding ensemble test. The average of 200 independently unfolded trials make up the unfolded spectrum. Left: 1 iteration. Right: 30 Iterations. The consistency in the central unfolded values are preserved over the number of iterations, while the edge bins diverge. The statistical uncertainty is seen to increase over the number of iterations.

is confounded of two factors: Firstly unfolded bin values move with the number of iterations. In the full unfolding space, some bins move closer to the truth while others move farther away with iterations. Overall, as can be seen in fig. 5.8 the ratio of unfolded to truth rises, however the uncertainty rises faster, bringing the test statistic into the plateau seen in fig. 5.9. While this yields a small test statistic, the ratio of unfolded to truth still leaves room for optimization, and will be discussed in more detail in chapter 8.

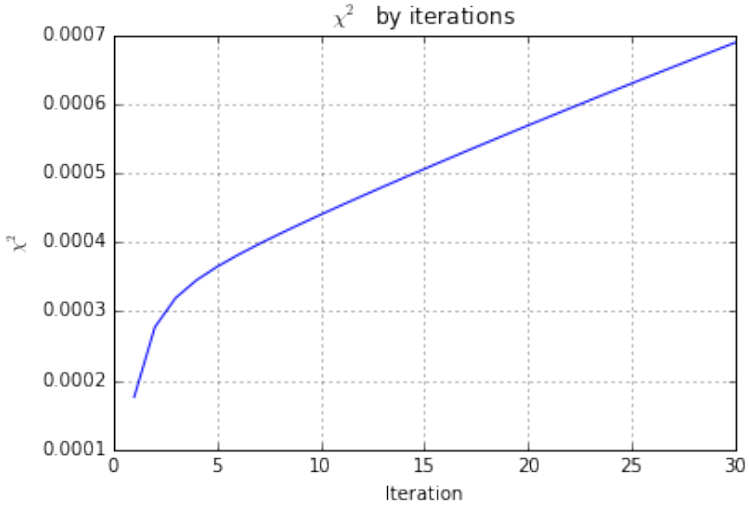


Figure 5.6: Unfolded to truth comparison in the Asimov case. The χ^2 is small to begin with as we expect, and increases as we update the prior and iterate. This is due to the unfolding already starting at the best possible match.

5.5 Unfolded Truth

An important aspect of any unfolding is the choice of truth variable - which value is unfolded. In essence: which result is the aim of the analysis? In this unfolding analysis the aim is maximally exploiting the model independence of the D'Agostini unfolding method. As such no assumptions regarding flux, oscillations or interaction cross sections are made. For the purpose of this analysis the truth side of the unfolding matrix must be constructed at such a point in the interaction chain, that those factors are part of the unfolded result.

$$N_{\mathrm{obs}}^{\mathrm{flav}} = \Phi \sigma N_{\mathrm{t}} T_{\mathrm{live}} \epsilon P_{\mathrm{osc}} \tag{5.6}$$

Here, the number of targets N_{t}, the livetime T_{live} and the efficiency ϵ are constants related to the experimental setup of IceCube, and those can be compensated for in the unfolding so that:

$$N_{\mathrm{unf}}^{\mathrm{flav}} = \frac{N_{\mathrm{obs}}^{\mathrm{flav}}}{\epsilon N_{\mathrm{t}} T_{\mathrm{live}}} = \Phi \sigma P_{\mathrm{osc}} \tag{5.7}$$

The unfolded variable is thus an event rate folded by the oscillation probability. This allows anyone using the unfolded data set to make their own assumptions regarding two of these variables, in order to measure the third one. While one can argue the oscillation parameters are well known, keeping the oscillation as part of the signal allows forward folding model tests of exotic oscillation scenarios using the data sample.

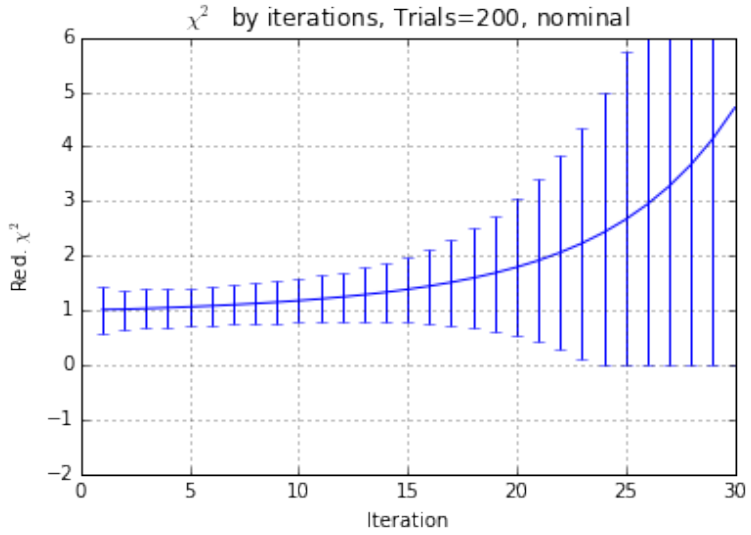


Figure 5.7: Average chi square comparison of 200 individually unfolded pseudo experiments to the MC truth. The unfolding now starts at a χ^2 of around one, which increases with the number of iterations. This is due to the pseudo experiments already being close to the prior knowledge of the unfolding machinery. We again essentially start at the best possible match between unfolding and MC truth. This plot illustrates the difference between the left and right panels of fig. 5.5.

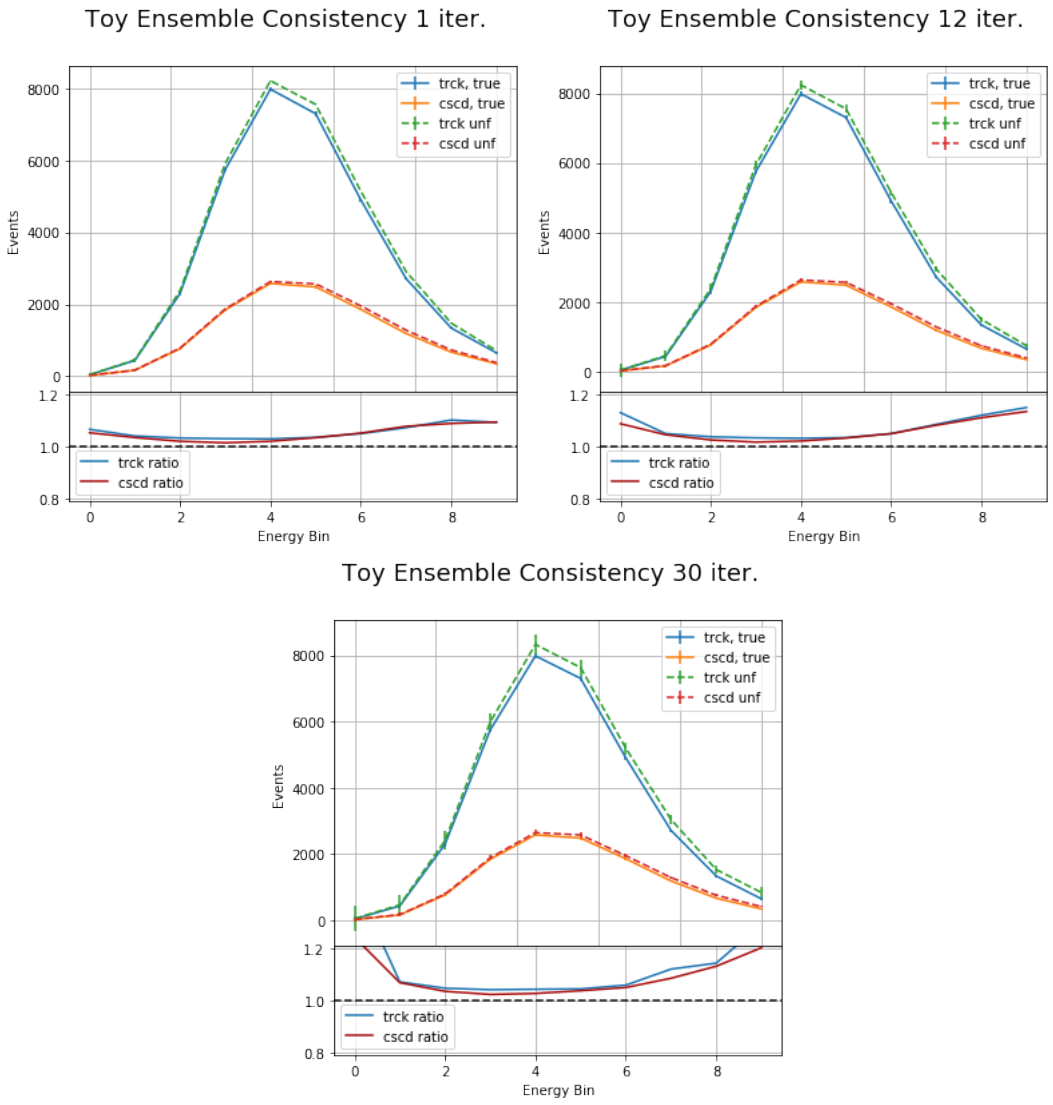


Figure 5.8: True and Unfolded spectra for the Toy model unfolding ensemble test. The average of 200 independently unfolded trials make up the unfolded spectrum. Top Left: 1 iteration. Top Right: 12 iterations. Bottom: 30 Iterations. Like in the previous case, the statistical uncertainty is seen to increase over the number of iterations.

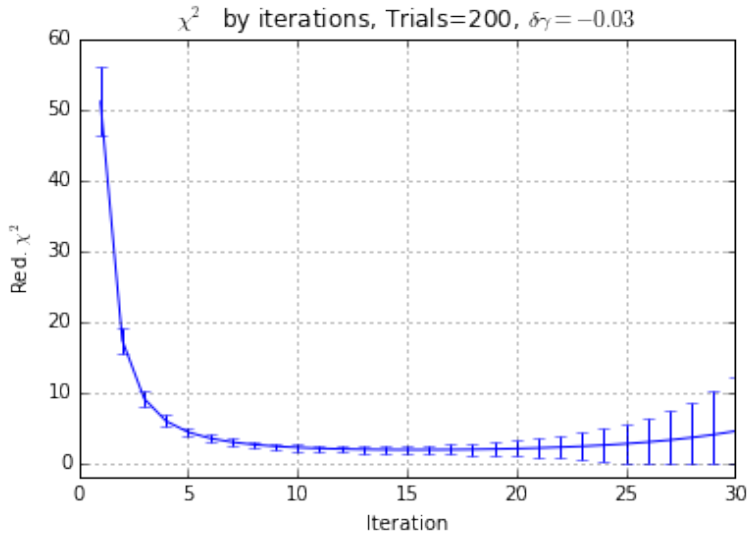


Figure 5.9: Average chi square comparison of 200 individually unfolded pseudo experiments to the MC truth. The pseudo experiments this time had their energy spectrum modified by a power law with index: $\delta\gamma = -0.03$. We here observe a rapid drop in the test statistic as the unfolding converges on the MC truth. A plateau is found around ~15 iterations, after which the statistical fluctuations get too much weight and again pull the unfolding away from the MC truth.

5.5.1 Truth Binning

The chosen binning has consequences for what information can be extracted in the analysis. Specifically, this approach to unfolding allows multi dimensional unfolding and events to be sorted in truth space based on their event classification. In this analysis we opt for ν_μ^{cc}, ν_{rest} and background, giving three bins in the '*Particle ID*' (PID) dimension. The choice of response matrix dimensionality and binning will be discussed further in section 8.1.

5.6 Unfolding the Event Rate

The response matrix relates number of true events to reconstructed events on a bin-by-bin basis. The aim of the unfolding is to measure the true event rate per volume ice. This can be achieved by modifying the prior used to construct the unfolding matrix. The MC event rate is weighted by the inverse number of targets, the inverse efficiency and the inverse livetime in order to reach the desired quantity, as described in eq. 5.7. The weighting is performed on a flavor basis since the individual flavors have different simulation volumes in the MC. The Muon background is likewise weighted based on the generation volume in the simulation. The background noise is always weighted by one, as the truth and reconstructed values are equivalent. When marginalizing to get the spectra in energy and $\cos(\theta_z)$ the energy spectrum is weighted by the bin center in order to better illustrate the shape of the spectrum. Finally, the bin widths are divided out and the units on the unfolded spectra in energy and $\cos(\theta_z)$ are $[\mathrm{m}^{-3}\mathrm{s}^{-1}]$ in both cases.

5.7 Unfolding the Flux

Much as the unfolding matrix can be weighted to unfold to an event rate per volume, it is also possible to unfold directly to a flux. The events are further weighted by inverse cross section, which yields a flux. However, care must be taken due to the uncertainty on the assumed cross sections, which can be large and lead to considerable shape variations in the final result.

6 Statistical Uncertainties

In the following we will discuss the unfolding analysis without systematics, in which we can then isolate the impact of the statistical uncertainties. The impact on uncertainty from having a finite simulation will inform our decisions on how many events it is necessary to generate, as well as how many dimensions can reasonably be used in the unfolding. In particular we investigate the impact of a finite data sample, as well as in our MC - specifically pertaining to the response matrix and the propagation of uncertainty through the unfolding steps. We start from an Asimov closure test and step by step relax the constraints in order to gauge the impact from each step.

6.1 Toy MC: From flux to event rate

In order to study the statistical impact of limited Monte Carlo simulation and limited data, we created a toy MC model to allow better understanding of the RooUnfold software before building the final analysis software. When using a toy MC there is access to the natural truth, which in our simulations then corresponds to the truth of nature. We simply choose functions to represent the expectation values for input data. All parameters are chosen to be in good agreement with our current knowledge.

6.2 Toy MC: True Smearing

The response matrix inherently represents the quality and properties of the detector and event reconstruction algorithms. In order to estimate the statistical uncertainty on the unfolding a response matrix with a limited number of entries must be generated. To that end, some discussion is necessary: in a real detection scenario the true spectrum is smeared according to the true response matrix of the detector, the knowledge of which will have to be estimated from Monte carol simulation. In this toy MC, the true detector smearing is defined from an n-dimensional Gaussian, where n is the number of dimensions used in the unfolding. This smearing function is then used to transform from true values to reconstructed values. Doing this then yields a toy MC Asimov data sample consisting of true- and reconstructed spectra as well as a response matrix. The toy MC thus provides us with the true smearing matrix via the analytical smearing function. A plot of the true response matrix in $\cos(\theta_z)$ and energy is shown in the top left of fig. 6.3

6.3 Asimov Closure Test

After generating the true event spectrum and smearing it to get the reconstructed event spectrum, the following closure test is performed. This will tell us two things: firstly whether the unfolding algorithm behaves as expected and gives meaningful results in a controlled test, and secondly what magnitude of a statistical uncertainty can be expected based on the event rates in the MC. Using the true smearing matrix as the input response matrix for the unfolding, the true event spectrum as the prior, and the reconstructed spectrum as our measurement, we perform the unfolding. If everything is in order we expect the unfolding to give an unfolded spectrum which matches the true event spectrum in just one iteration. The statistical error bars in this case represent the 68% interval in which data from unfolded realizations of the reconstructed spectrum should fall. The result of the Asimov closure test is shown in fig. 6.1 and with a $\chi^2 \sim 10^{-17}$ the results are seen to be consistent.

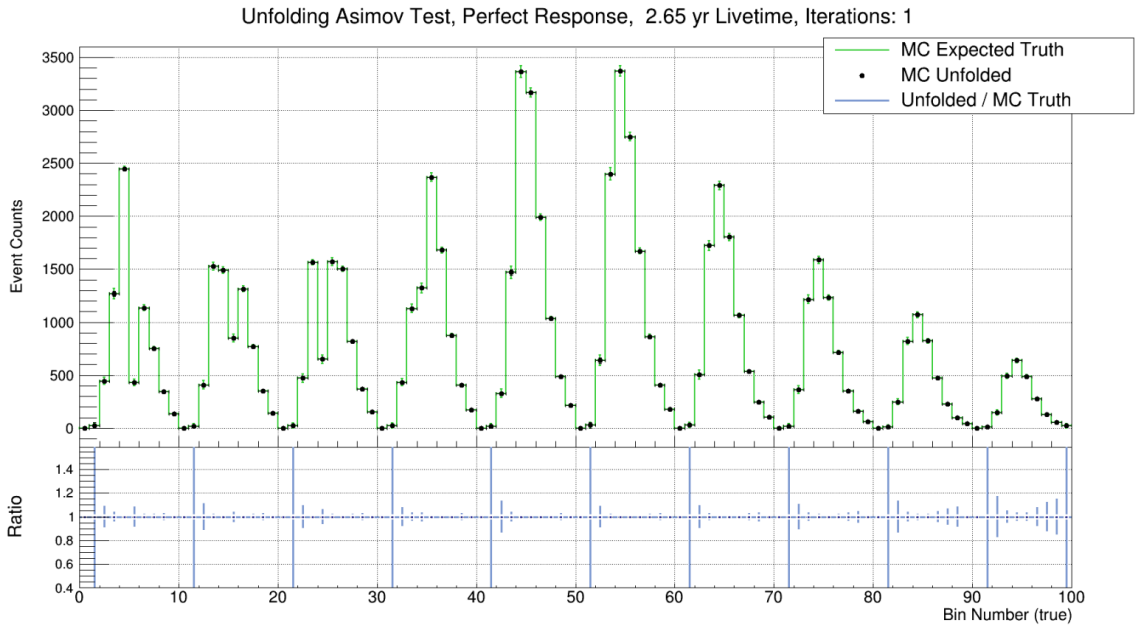


Figure 6.1: Unfolded and true distributions for the Asimov unfolding test, flattened from 2D, so every ten bins constitutes an energy slice in $\cos(\theta_z)$ space. Unfolding converges to the true value in just one iteration and statistical errors are below 10% except for bins with very low statistics.

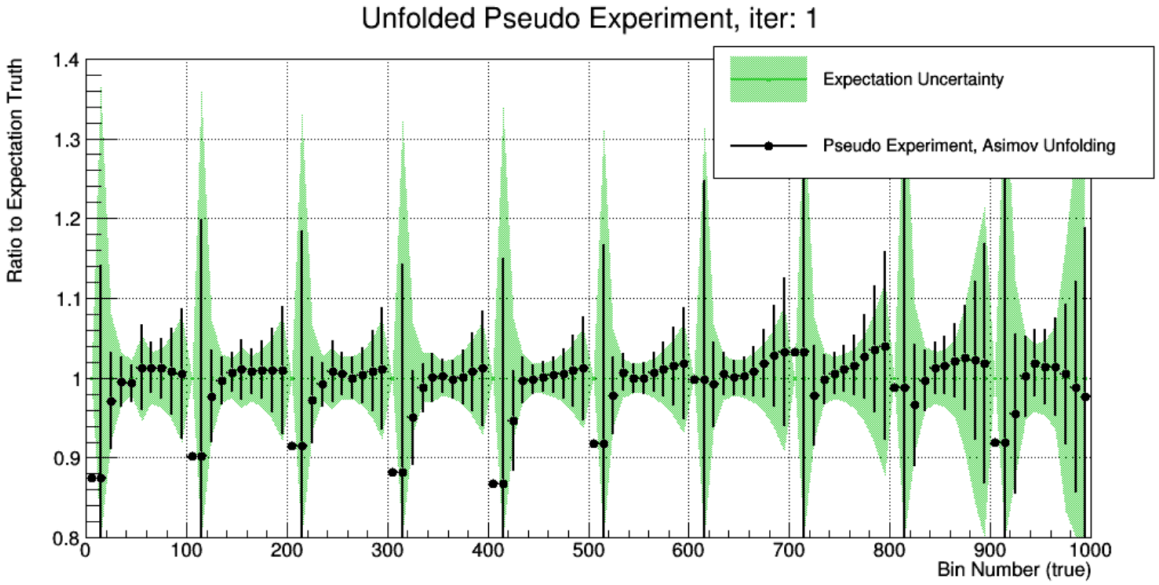


Figure 6.2: Ratio of unfolded to truth, for the unfolded pseudo experiment. The green shaded bands represent the uncertainty range from the Asimov case, in which the unfolded pseudo experiment should fall 68% of the time. The unfolding converges in one iteration and all unfolded data points fall within the expected region. The black error bars represent the 1 sigma uncertainty range of the underlying true distribution.

6.4 Limited Data Statistics

Proceeding from the Asimov unfolding closure test in the previous section, a pseudo experiment is drawn from a Poisson distribution around the expectation value of the reconstructed spectrum. This is then unfolded using the true smearing matrix. The results give a concrete example of what a realization of the toy MC truth could be. The unfolded data points are expected to fall within the one sigma green shaded area 68% of the time. The unfolding and the Asimov closure test are shown in figure. 6.2 and show good agreement between Asimov and pseudo experiment closure tests.

6.5 Propagation of Statistical Uncertainties in the Response Matrix

There is a known statistical uncertainty on the response matrix as well as the input pseudo experiment. To calculate the statistical uncertainty on the unfolding, the statistical uncertainty on the unfolding matrix must be calculated and correctly propagated through the unfolding iterations, as the unfolding matrix is applied to the input pseudo experiment. This can be expressed as an error propagation matrix:

$$\frac{\partial \hat{n}(C_i)}{\partial n(E_j)} = M_{ij} + \sum_{k=1}^{n_E} M_{ik} n(E_k) \left(\frac{1}{n_0(C_i)} \frac{\partial n_0(C_i)}{\partial n(E_j)} - \sum_{l=1}^{N_C} \frac{\epsilon_l}{n_0(C_l)} \frac{\partial n_0(C_l)}{\partial n(E_j)} M_{lk} \right) \tag{6.1}$$

Here M is the unfolding matrix, $\hat{n}(C)$ are the estimated causes, ie: the unfolded values and $n(E)$ are

the measured effects, in this case the pseudo experiment or actual data. ϵ is the efficiency and the index on n illustrates the number of iterations, while indices i, j, k and l represent the bin numbers. This expression then depends on $\frac{\partial n_0(C_i)}{\partial n(E_j)}$, which is nothing but $\frac{\partial \hat{n}(C_i)}{\partial n(E_j)}$ from the previous iteration [72] .

6.6 Limited Knowledge of the Detector Response

Unfortunately, the process for generating the response matrix is not quite as simple as was the case for the true smearing matrix. In reality the response matrix is estimated from event-by-event reconstructed MC, leaving us with an approximation to the true response matrix of IceCube, as depending on our MC statistics. The more statistics the better. In order to then generate a fair estimate of our response matrix we must take into account our limited MC statistics. With the true smearing matrix and our pseudo experiment well defined, we generate different realizations of the detector response matrix by drawing samples from the true smearing matrix. Shown in fig. 6.3 are the true smearing matrix and three response matrices generated by 20000, 200000 and 500000 samples respectively. We then independently unfold the pseudo experiment from the previous section three times using these statistically different response matrices. Any difference between the pseudo experiment closure test and the tests performed here necessarily arise from the limited statistics of the response matrix used. As we proceed, once we have enough statistics in the response matrix, we expect the unfolded data points to fall within the one sigma shaded region 68% of the time and have error bars similar to those of the Asimov closure test. The results are shown in fig. 6.4, and we qualitatively conclude that we need on the order of 500000 MC events to reach the precision of the Asimov closure test. The number of events included in the MC sample is roughly 491 thousand as described in section 5.3, which is sufficient for the purpose of the unfolding.

6.7 Binning Optimization

In order to get a reasonable statistical uncertainty in our final sample, and across all bins, we optimize the bin size based on event counts in each individual bin. This is done by simple numerical integration in both the $\cos(\theta_z)$ and energy dimension individually. The binning is optimized to give a statistical uncertainty on the event rate of about one percent in each bin. The resulting binning is twelve irregularly shaped bins in both $\cos(\theta)$ and energy, which will be used henceforth. The bin edges are shown in table 6.1. Although this binning has been optimized in regards to the statistical uncertainty, later investigations of the sample coverage proved it necessary to further expand the binning. This is discussed in section 8.14.1, and the updated binning is listed in table 6.2. This updated binning is utilized several places in the following.

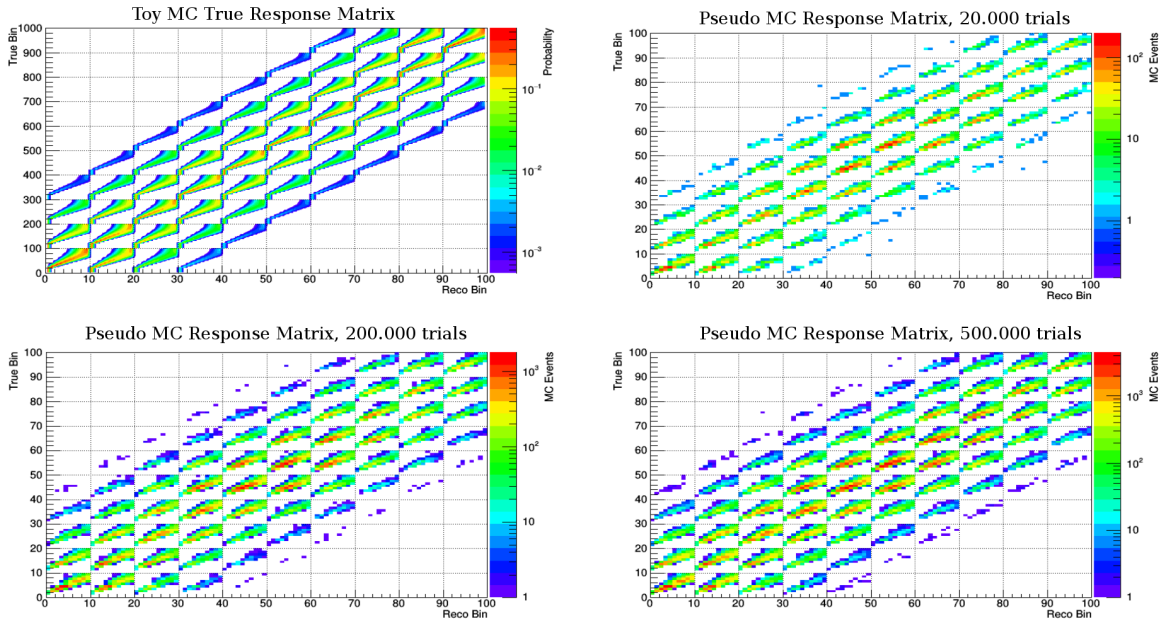


Figure 6.3: Toy MC Response Matrices. Top left: true smearing/response matrix used in the toy MC. Other panels: realizations of the response matrix for different sample sizes, known as pseudo MC response matrices. Top right: 20000 samples. Bottom left: 200000 samples. Bottom right: 500000 samples. With increasing number of samples the pseudo matrices converge on the shape of the true smearing matrix.

True	Energy [GeV]	1.0	6.04	8.44	10.72	13.36	16.72	21.04
	$\cos(\theta_z)$	-1.0	-0.81	-0.62	-0.44	-0.27	-0.14	-0.05
	Energy [GeV]	26.68	34.11	44.91	63.87	109.28	600.0	
	$\cos(\theta_z)$	0.04	0.13	0.25	0.41	0.64	1.0	
Reconstructed	Energy [GeV]	1.0	6.98	9.73	12.21	15.14	18.43	22.38
	$\cos(\theta_z)$	-1.0	-0.82	-0.65	-0.49	-0.35	-0.22	-0.09
	Energy [GeV]	27.49	34.22	43.58	57.99	86.75	300.0	
	$\cos(\theta_z)$	0.03	0.16	0.31	0.48	0.70	1.0	

Table 6.1: Binning in true and reconstructed space. Bin edges are set by numerical integration of the upper bin edges, and are independent of the detector resolution.

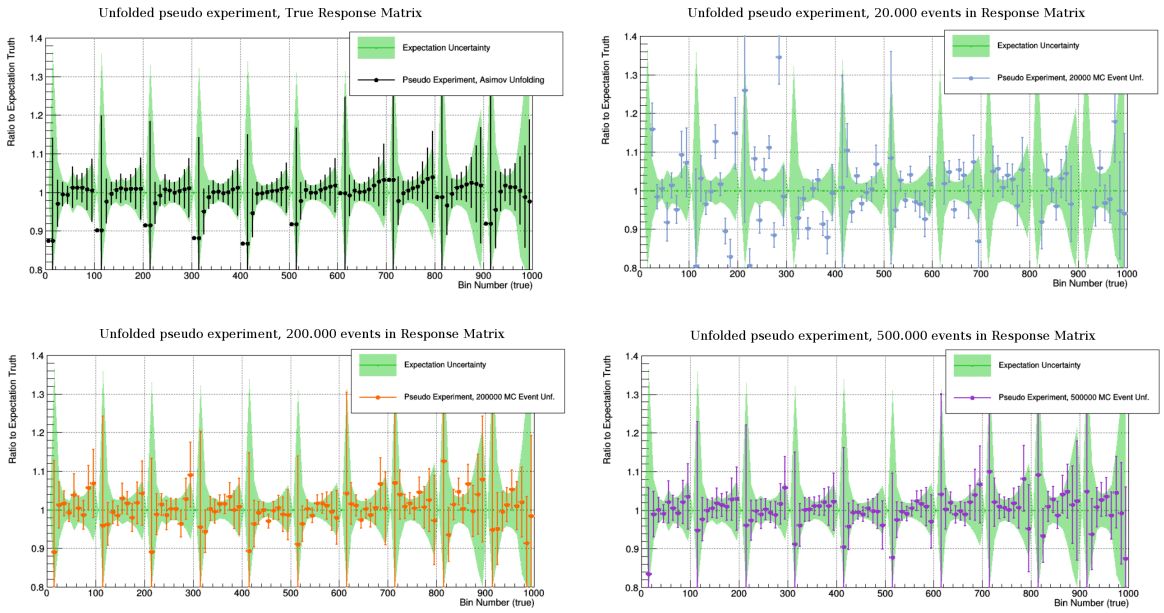


Figure 6.4: Results from the toy MC unfolding when using different response matrices, each generated from a separate sampling of the true smearing matrix. The green band represents the expectation uncertainty into which 68% of all samples are expected to fall. As the number of trials in the response matrix increases, the unfolding converges to the true smearing matrix case. The qualitative conclusion is that we need on the order of 500000 MC events to reach the precision of the Asimov closure test.

True	Energy [GeV]	1.0	6.04	8.44	10.72	13.36	16.72	21.04
	$\cos(\theta_z)$	-1.0	-0.81	-0.62	-0.44	-0.27	-0.14	-0.05
	Energy [GeV]	26.68	34.11	44.91	63.87	109.28	200.0	1000
	$\cos(\theta_z)$	0.04	0.13	0.25	0.41	0.64	1.0	
Reconstructed	Energy [GeV]	1.0	6.98	9.73	12.21	15.14	18.43	22.38
	$\cos(\theta_z)$	-1.0	-0.82	-0.65	-0.49	-0.35	-0.22	-0.09
	Energy [GeV]	27.49	34.22	43.58	57.99	86.75	300.0	1000
	$\cos(\theta_z)$	0.03	0.16	0.31	0.48	0.70	1.0	

Table 6.2: Analysis binning as updated due to investigation in section 8.14.1. The previous last bin edge has been altered to be at 200 GeV, and another bin has been added from 200 GeV to 1 TeV.

7 Systematic Uncertainties

The systematic uncertainties cover any uncertainty related to the detector, simulation or other effects not arising from the statistical nature of the sample. Since the binning has been selected to have a relative statistical uncertainty on the signal of maximum 1% in each bin, the measurement precision is likely systematics dominated and careful consideration must be made for each systematic effect. The systematic uncertainties are divided into two broad categories, by their method of implementation: 'Continuous Systematics' and 'Discrete Systematics'. The systematic effects are implemented in the simulation chain on an event-by-event basis as a weight to each individual event. In the case of i different systematic uncertainties, the final weight of each event in the simulation can thus be expressed as:

$$w = \prod_i \delta w_i, \tag{7.1}$$

where $\delta w_i = 1$ is the nominal case of the systematic parameter in question (corresponding to the standard simulation as discussed in chapter 4). This section will discuss the implementation of each individual systematic and the impacts on the analysis will be discussed in chapter 8.

7.1 Systematics Strategy

In order to estimate the impact from systematics on the final result, two main methods have been employed: a bracket approach and a Gaussian sample approach. Here will be presented the considerations that led to the final choice of implementation. The goal of any systematics implementation is to estimate how much any measurement will differ from the underlying truth, due to systematic effects. The central limit theorem [78] predicts that the mean measured value of an experiment will tend towards the underlying truth as the number of samples increases. Thus, statistical uncertainties can be alleviated by increasing the measurement statistics in the experiment. For systematic uncertainties there is no such luxury because any uncertainty in the measuring apparatus can move the central value of the measurement away from the underlying truth. Unlike a known bias, which can be corrected for, the exact scale of the impact from systematic uncertainties is generally unknown. Thus, the quest of any systematic uncertainties study, must be a quantization of this effect.

Furthermore there is the question of how to implement the systematics in the analysis ansatz: as both a signal and unfolding matrix are estimated from MC, two options present themselves. Previous unfolding analyses [79] had implemented the systematics on the input spectrum and kept the knowledge of the detector smearing constant. It may seem counter intuitive to modify the measured spectrum to account for systematic effects, but it does yield a representation of how the event rates would respond to changes in detector properties. However, such a method would not fully cover the effect any systematic would have on the detection efficiency. A different, and more traditional approach is to implement the systematic effect on the MC spectrum and the detector smearing while keeping the input constant. In principle, for an unfolding analysis like the one presented in this thesis, it would be easy to imagine the two methods to be equal, since what should a-priori matter is the difference between the input spectrum and the prior used for constructing the unfolding matrix. However, more careful analysis performed in the early stages of this work, revealed the two methods do not yield completely identical results, partly due to the relative spectrum shifts combined with the statistical impact on the unfolding matrix.

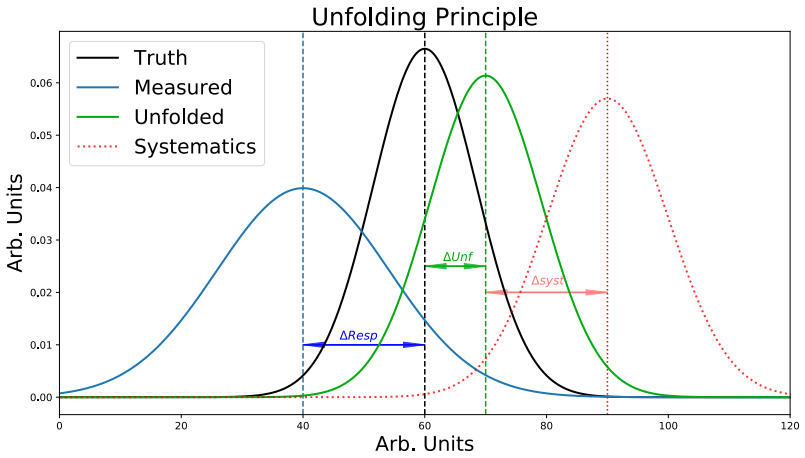


Figure 7.1: Representative plot of the unfolding procedure including systematic uncertainties as applied in this analysis. The underlying truth of nature is shown in black. The blue curve represents the measured signal, which would usually be subject to statistical- and systematic uncertainties in a standard measurement. The green curve is the unfolded distribution - the estimate of the underlying truth. The systematics impact on the unfolded measurement is represented in red.

In this work, the decision was made to keep the input spectrum constant and to implement the systematics on the MC and detector response. This means the input measurement is kept as-is, and the method will give an accurate estimate of the systematic uncertainty in the unfolded measurement. When building the unfolding matrix both the true and reconstructed values of the input MC are used, so it is also necessary to consider the impact of systematics on these (see sec. 5.2). In turn, this impact leads to a shift in the unfolding matrix, which again leads to a systematics shift in the unfolded spectrum. The principle of the systematics procedure is illustrated in fig. 7.1, where the initial shift from the true distribution to the measured one, ΔResp, is due to the response matrix, and the ΔUnf arises from the unfolding error, as described in section. 5.2. Statistical and systematic uncertainties will manifest in addition to this unfolding error, and due to the true value of nature being unknown will be implemented on the unfolded spectrum as illustrated. In the following the different systematics will be described along with their impact on the event rates in both true and reconstructed space.

7.2 Systematics Chain

We need to consider any systematic arising from experimental effects which are not directly related to the quantity we want to measure. As discussed in section 1 the main result of this analysis is an as model-agnostic as possible unfolding of the atmospheric neutrino event rate by detector volume. Let us briefly consider the physics of the analysis chain: A cosmic ray primary interacts in the atmosphere, generating daughter particles, eventually leading to neutrino emission. This relationship is governed by the cascade equation as discussed in section 2.4. The exact relationship is uncertain, but can to some extent be tuned by correlation with the Muon flux. The generated neutrinos propagate through the atmosphere and the Earth and upon arrival at IceCube they will have undergone neutrino oscillation

of varying baselines. This yields a dependence on oscillation parameters and matter interactions. Only a certain fraction of neutrinos arriving will interact in the ice, introducing uncertainties on the interaction cross section. The interaction yields a particle shower in the IceCube fiducial volume (in the case of a charged current interaction also with an outgoing Cherenkov track), with an electromagnetic component. The expanding cascade and outgoing Cherenkov track are well understood, have been implemented in detailed simulations, and carry little to no uncertainty. The photonic output signal on the other hand is influenced by the optical properties of the ice, which are not nearly as well known and must be taken into account as a systematic effect. Eventually photons reach the string bore holes, and some are detected at the IceCube DOMs. The bore holes have different optical properties to the bulk ice and the DOMs are not 100% efficient - both are effects which must be included as systematic uncertainties. The combined effect of the systematic uncertainties on the event selection efficiency must also be included. While the charge and timing information collected from the PMTs has a high level of precision, the derived quantities of energy and incoming neutrino direction become uncertain due to the quality of the reconstruction algorithm. This results in an amount of detector smearing in those two quantities, which is described by the detector response matrix. The unfolding procedure is able to overcome this smearing by reversing the effect via the unfolding matrix. Unlike the response matrix, the unfolding matrix depends on the input prior, and is thus affected by systematic uncertainties. In summary, each step in the chain of physical processes leading to neutrino detection is associated with specific systematic uncertainties:

- Neutrino flux: flavor ratios, neutrino-anti-neutrino ratios, spectral index
- Oscillations: oscillation parameters, matter effects
- Interaction: cross sections
- Ice Properties: bulk ice scattering and absorption. Hole ice properties.
- Photon detection: DOM efficiency (which impacts triggers, veto efficiency, reconstruction precision)
- Detector response (overcome by the unfolding)

As such we see that certain systematics in the analysis ansatz are associated with the unfolded quantity, while others must be accounted for in the unfolding analysis. Because we unfold the event rate by volume, the measurement should be independent of any systematic related to the flux, oscillations or cross sections. Although these measures are introduced in the prior term, the unfolding is able to overcome that bias, given optimization of the number of iterations. On the other hand, several systematic parameters require careful consideration as part of the analysis. They encompass the uncertainties related to the bulk ice properties, the hole ice properties and the efficiency of the DOMs. The detector response is estimated from MC depending on these quantities leading to an impact on it from the systematics. The systematics of interest have been listed in table 7.1 and are discussed in more detail in the following sections.

Discrete Systematics		
Parameter	Value	Prior
DOM Efficiency	1.0	$\pm 10\%$
Hole Ice, p	25	± 5
Bulk Ice Scattering	1.0	10%
Bulk Ice Absorption	1.0	10%

Continuous Systematics		
Parameter	Value	Prior
Livetime	$4.8yr$	1%
Muon Normalization	1.0	10%
Noise Normalization	1.0	10%

Table 7.1: Systematic uncertainties under consideration. The parameters, their central values as well as our prior assumptions are listed. Top: Systematics implemented in a re-weighting scheme. Bottom: Systematics implemented in a discrete systematics scheme. Where possible central values and priors have been taken from [50].

Systematic	MC Simulated Values						
DOM Efficiency	0.88	0.94	0.97	1.00	1.03	1.06	1.12
Hole Ice		0.15	0.20	0.25	0.30	0.35	
Bulk Ice Scattering				1.0	1.1		
Bulk Ice Absorption				1.0	1.1		

Table 7.2: Discrete Systematics included and the values at which they are sampled.

7.3 Discrete Systematics

Certain parameters are known as 'discrete systematics', and have an effect on low-level IceCube simulation, such as affecting the trigger conditions or the photon propagation in the ice. Changing these parameters requires changing the systematic in question in the icecube simulation chain, and then re-processing the data. The GRECO processing chain and reconstruction is used in order to obtain the new estimate. As this process is very computationally intensive the systematics are sampled at discrete values (hence the name), and there has been considerable effort in IceCube to produce low-level standard simulation sets, which are then processed through the GRECO data selection. The analytical form of systematics is then discretized by the binning in the following way: a fit is performed to the bin-by-bin event rate at L7, yielding back the functional form representing the evolution of the event rate as it develops as a function of the discrete systematic, in that individual bin. This functional form for each bin is then sampled on an event-by-event basis to yield the correct weight for each event. In summary: using these fits it is possible to set the discrete systematic parameter to any desired value within the analysis framework and simply weight the nominal event rate based on the relative strength of the fit. Several discrete systematics are implemented separately for neutrinos and muons, as they are simulated separately. The implementation has been developed, tested, and verified independently on the two sets. In the sample configuration the discrete systematics for both neutrinos and muons are kept to the same level when combining the simulations, in order to preserve consistency and physical correctness.

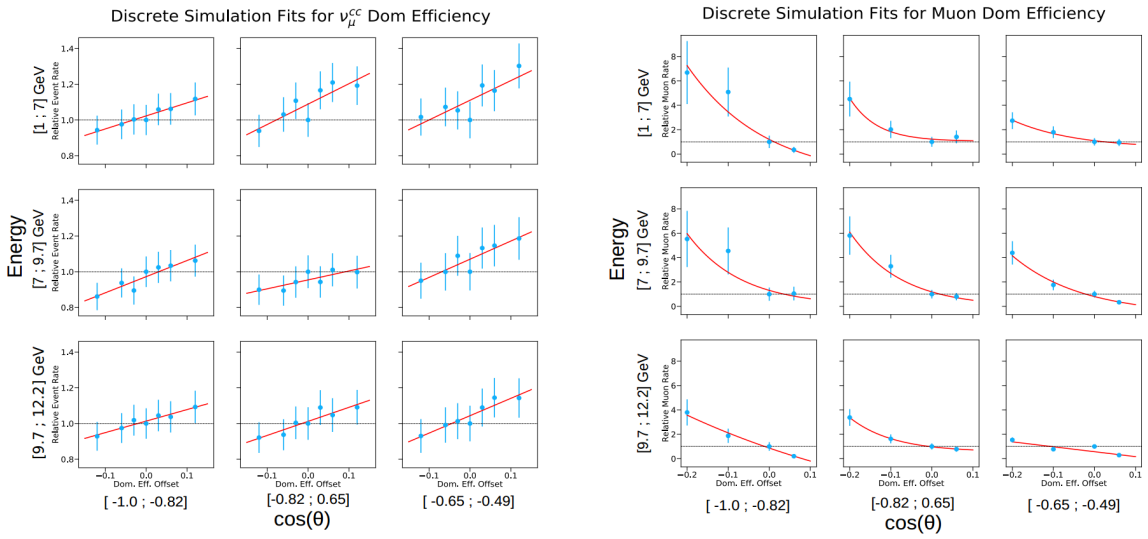


Figure 7.2: Fits to the discrete simulation sets for the event rate as a function of DOM efficiency. Left: neutrinos. Right: muons. Blue shows the event rate relative to the nominal data set, while red indicates the fitted function. The black dashed line indicates the nominal value of one. For neutrinos the event rate relationship with the DOM efficiency is linear. For Muons the functional form generally follows an exponential decay, as decreasing the DOM efficiency exponentially increases the muon background. However, a few bins display more of a linear relationship, which is also well described by the exponential fit. Shown here are the fits for nine bins.

7.3.1 DOM Efficiency

The DOM efficiency describes the probability for the optical modules to detect an incoming photon and is an estimate of the combined efficiency of the quantum efficiency of the PMT itself, as well as the transmission coefficients of the glass and the optical gel. It is usually listed relative to the expected PMT-only efficiency. Ie: when the DOM efficiency is 100% the DOM is equally efficient to a standard PMT.

The DOM efficiency was measured prior to deployment using a sample of 16 DOMs, and was found to have a relative uncertainty of 7.7% [51]. There are ongoing efforts in IceCube to improve the knowledge of the DOM efficiency, however, in this work a conservative approach will be taken and the DOM efficiency is set to $1.00 \pm 10\%$. Like other discrete systematics the DOM efficiency is fitted as a discrete systematic with a linear functional form for each analysis bin, giving the relative weight in the individual bin as a function of DOM efficiency. A selection of the fits are shown in the left panel of fig. 7.2. All analysis bins follow a linear relationship, however, for some bins the scale is different from what is expected with a change in DOM efficiency having symmetric effects regardless of whether the DOM efficiency parameter is raised or lowered. In these cases the fit takes precedent, as our standard simulation value could be different from the truth and the weight of the events falling in that particular bin is determined by the value of the fitted function. A higher DOM efficiency generally leads to more detected photons at each DOM. More detected light in turn leads to more triggers and

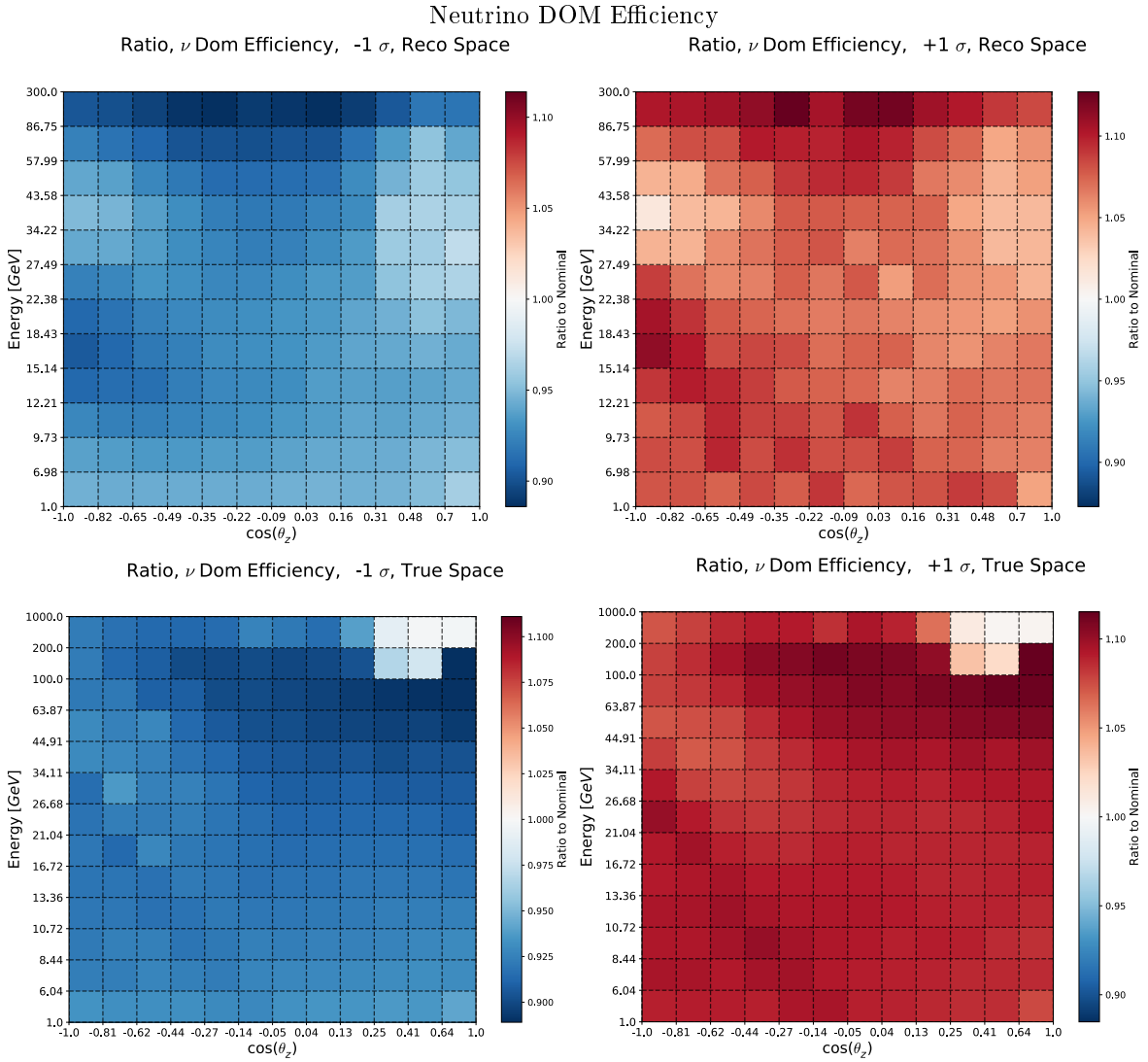


Figure 7.3: Neutrino DOM efficiency systematic impact on the final simulated spectrum including muon background and noise. Shown is the 2D event rate at the nominal $\pm 1\sigma$ values. Upper: reconstructed space. Bottom: true space. The maximum and minimum event rates change by approximately 10 as compared to the nominal value, which is what is to be expected when changing the DOM efficiency by 10%. Note the effect is less pronounced in the high energy downgoing bins, as the majority of the signal in these bins is made up of atmospheric muons.

yields better information of interactions in the ice. This gives a more precise event reconstruction, eventually leading to a higher neutrino event rate at final analysis level. The effect of varying the DOM efficiency to ±10% on the sample at final analysis level is shown in fig. 7.3

7.3.2 Muon Dom Efficiency

The impact of the DOM efficiency on the background muon rate is estimated in the same way as other discrete systematics: a fit across the simulation sets is performed for each bin with events falling in that bin, getting re-weighted based on the desired value of the systematic. Note however, that the DOM efficiency has an inverted impact for the muon background: increasing the DOM efficiency increases the efficiency of the muon veto yielding a lower event rate, while decreasing the DOM efficiency exponentially increases the number of background muons because the veto system suffers. This effect is non-linear and the correlation is modeled as an exponential decay, which is implemented in the functional form of the bin-by-bin fit. For the vast majority of bins this relationship holds true, however, a few bins display a more linear relationship, which can still be well described by the exponential form, as seen in fig. 7.2. The impact from combining the asymmetric effects from both the neutrino- and muon DOM efficiencies on the final spectrum, are shown in fig. 7.5.

7.3.3 Hole Ice

The drilling process for installing the IceCube strings involves melting the packed Antarctic ice. As this water re-freezes over a period of days, impurities and gases are pushed to the center of the drill hole, leading to a vertical so-called 'bubble column' with a diameter of about $16cm$ located at the center of the drill hole. Unfortunately this bubble column has a much shorter scattering length than the surrounding ice, yielding the unfortunate effect of scattering light which would have otherwise reached the DOMs. Thus, this effect is modeled as a correction to the DOM angular acceptance, which is applied in the simulation when a photon is incident on a DOM. Many initiatives have been taken to investigate the bubble column, the first can be attributed to IceCube's predecessor Amanda [73], with the latest being implemented in the future IceCube upgrade. Here the discussion will focus mainly on results from studies using the flashers on the DOMs. The acceptance curve is modeled as follows:

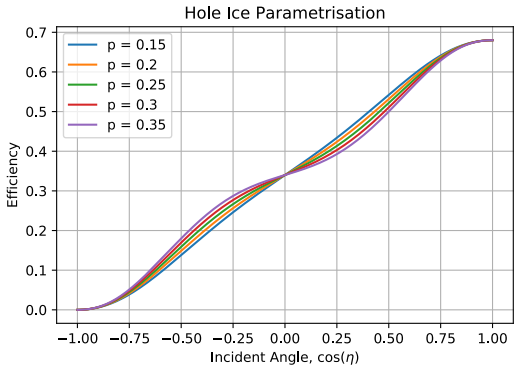


Figure 7.6: Angular acceptance of DOMs due to the effect of the hole ice.

$$0.34\left[1+1.5\cos(\eta)-\frac{\cos^3(\eta)}{2}\right]+p\cos(\eta)(\cos^2(\eta)-1)^3 \tag{7.2}$$

where η is the angle of acceptance and p is a discrete constant, with a nominal value of 0.25. This function is illustrated in fig. 7.6 showing each of the discrete values at which the systematic is sampled (corresponding to the values in table 7.2). Based on this functional form, an event rate dependence

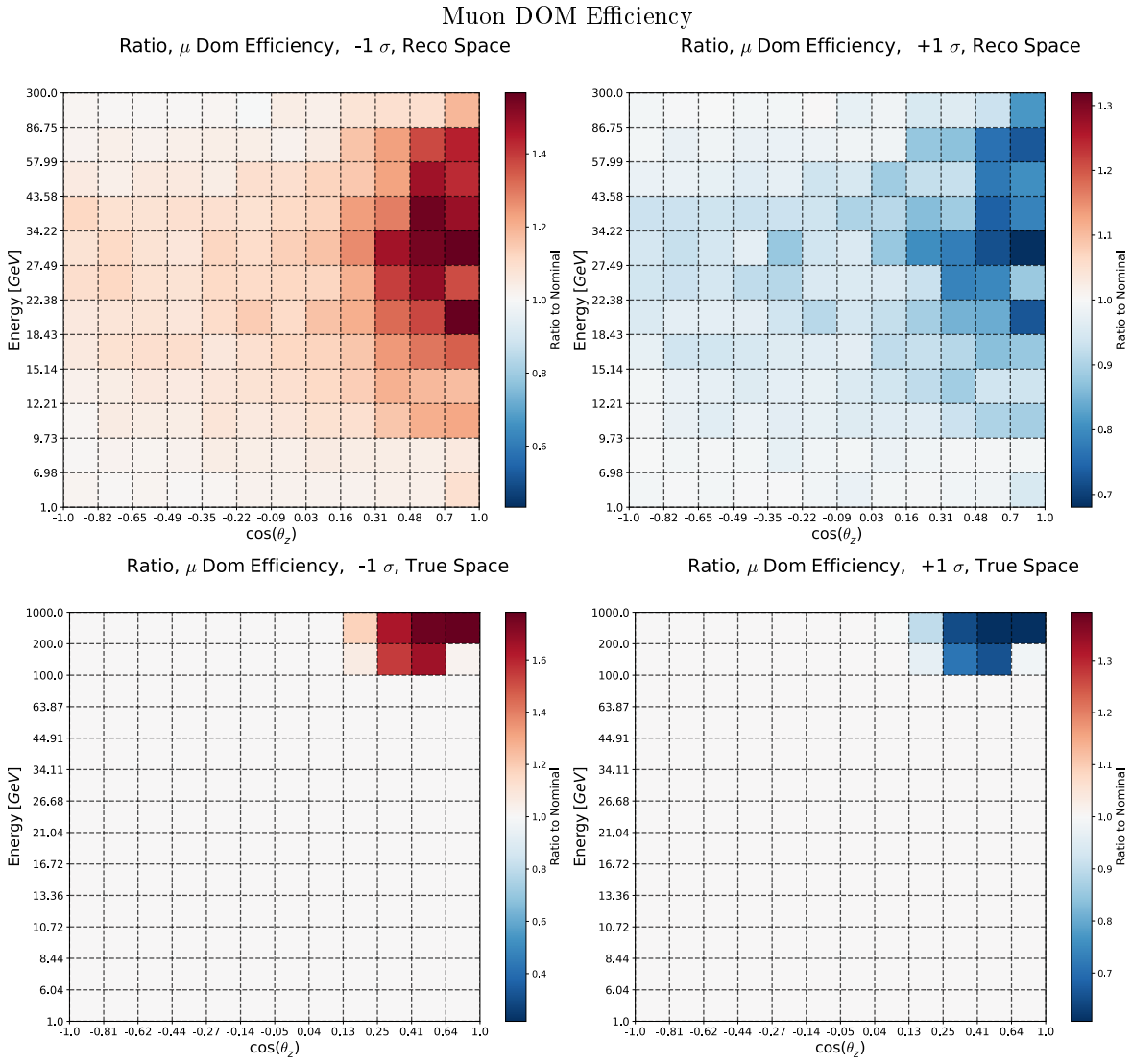


Figure 7.4: DOM efficiency impact implemented only on the muon part of the simulation. Top: reconstructed space. Bottom: true space. Notice the asymmetry arising from the exponential decay behavior of the muon veto: in the case of an increase in DOM efficiency we see hardly any decrease in the muon rate, however, with a decrease in DOM efficiency the muon rate rises exponentially.

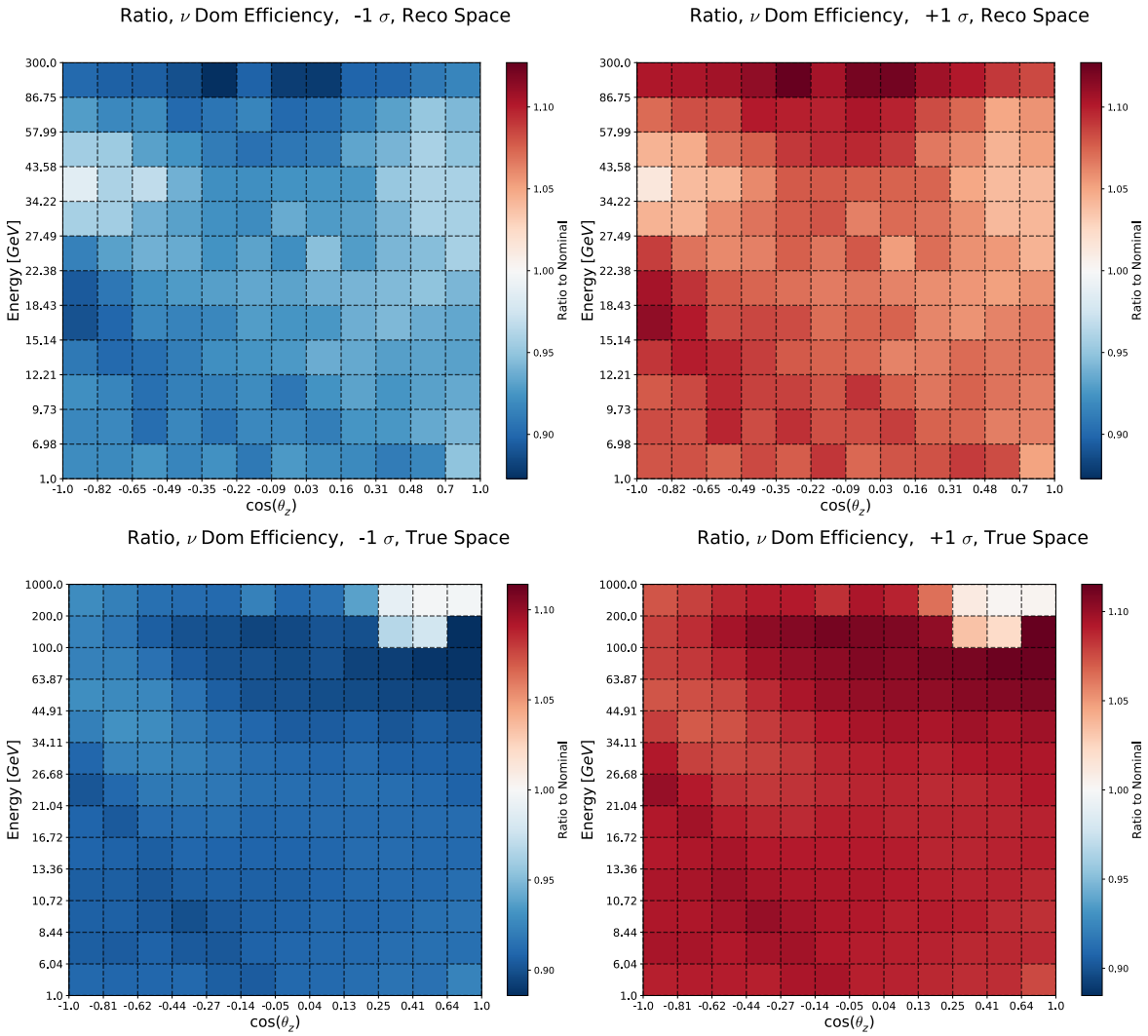


Figure 7.5: Combined DOM efficiency impact on the final simulated spectrum including muon background and noise. Top: reconstructed space. Bottom: true space. Notice the clear asymmetric effect arising from the exponential decay behaviour of the muon veto - the regions dominated by neutrinos lose events in case of a DOM efficiency reduction, while the regions dominated by muons gain events.

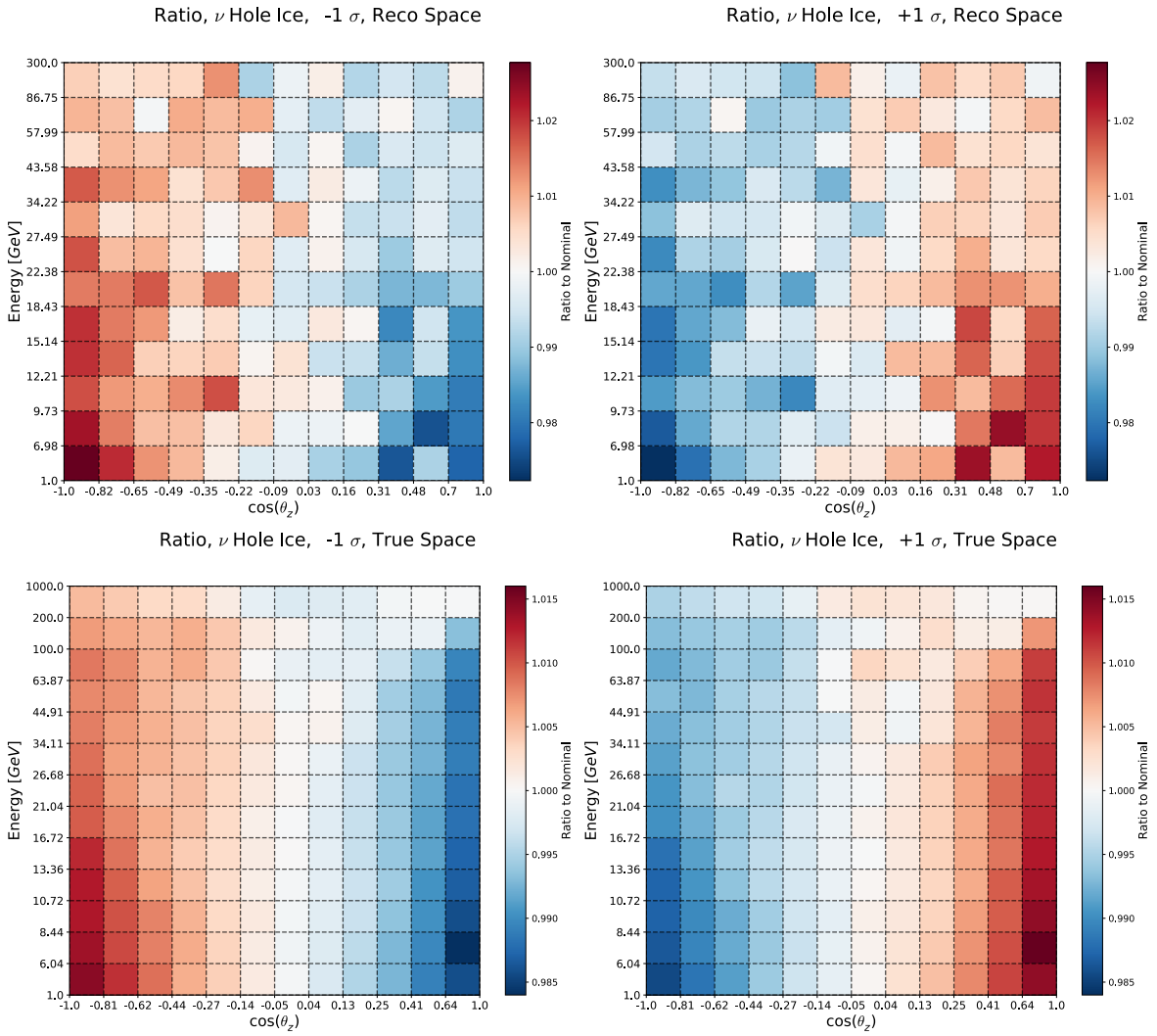


Figure 7.7: Effect of implementing the hole ice on the final level event rates. Upper: reconstructed space. Lower: true space. The effect depends on arrival direction, this is true in reconstructed space, but the same effect is even more clearly evident in true space.

on arrival direction is expected when implementing the systematic, and this is exactly what is seen in fig. 7.7.

7.3.4 Bulk Ice

The bulk ice discrete systematic is implemented as a modification of two separate physical effects: the scattering length in the ice and the absorption length. With a reduced scattering length, less light is scattered during photon propagation and it is reasonable to expect the reconstruction algorithm to perform better both for neutrinos and muons. An effect not unlike the DOM efficiency is thus expected: a reduction in scattering length is expected to lead to more neutrino events and fewer muon events, due to better veto performance. Note that the parameters are implemented in such a way that negative values of σ implies *less* scattering/absorption corresponding to a longer scattering/absorption length and vice versa. It is thus a priori expected that an increase in either of these parameters will lead to a decrease of the observed event rate. The bulk ice is modeled in layers of 2cm width in which the properties are roughly identical. Furthermore, each of these layers have a preferred direction of light propagation along which there is less scattering/absorption. The discrete sets used here vary only the scattering/absorption, not any of the other ice properties, as simulations sets including these were not available at the time of writing. Longer term, these simulation sets should be produced and included in future analyses.

7.4 Continuous Systematics

Unlike the discrete systematics, the continuous systematics are described by continuous parameters in the MC software, as given by eq. 7.1. A short overview of the three continuous systematics included will be given in the remainder of this subsection.

7.4.1 Livetime

The total time of active measurement using the IceCube/DeepCore detector is known as the livetime. The livetime is calculated based on the timestamps from the IceCube good run list, for the runs included in the GRECO sample. The livetime is known to second precision for each run, and is summed for all runs. The livetime is introduced as a constant scaling parameter over all the parameter space, meaning event rates scale linearly with livetime. Noise rates can influence the effective livetime due to electronics dead time, which may lead to the effective livetime being reduced, and with livetimes on the scale of 5 years or more the solar cycle can change the number of neutrino events due to changes in the primary cosmic ray flux. These extra considerations are effectively included in the livetime systematic parameters, the prior for which is conservatively set to one percent.

7.4.2 Muon Scale

A scaling factor is introduced to adjust the level of background muons introduced in the MC described in sec. 4.4. While in reality the muon rate and the neutrino rates are correlated, in the IceCube MC, they are independent parameters. The muon rates have previously been measured and an uncertainty of five percent has been assigned to the muon rate in this work. While the spectrum shape might differ, early investigations in this work revealed a change in the muon spectral index to be of negligible importance in relation to this analysis. The relative impact on the reconstructed event rates are shown in fig. 7.10, and show an increase in event rate up to about 6% in the high energy downgoing region.

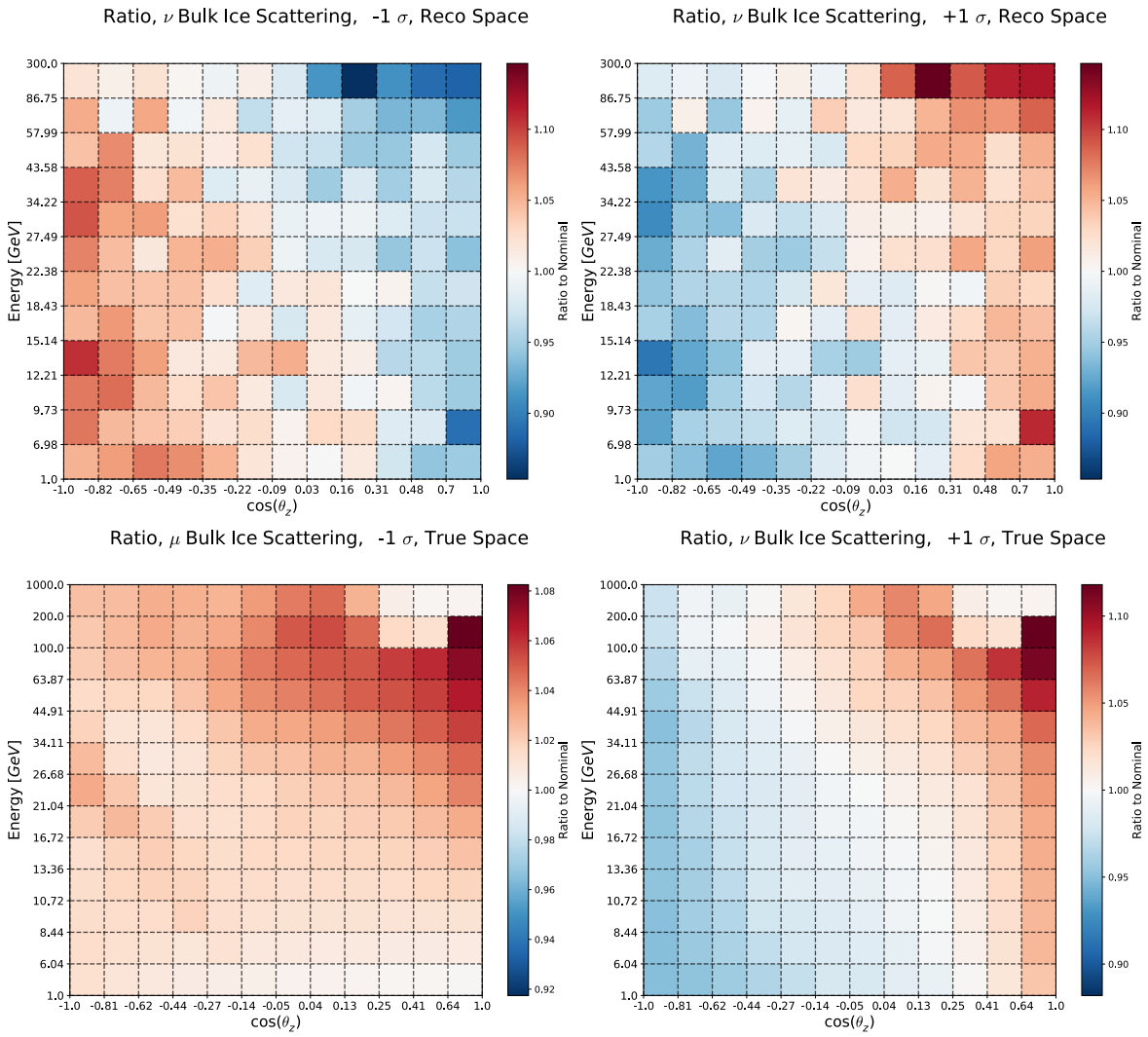


Figure 7.8: Effect of the bulk ice scattering length systematic implementation on the final level event rates. Upper: reconstructed space. Lower: true space.

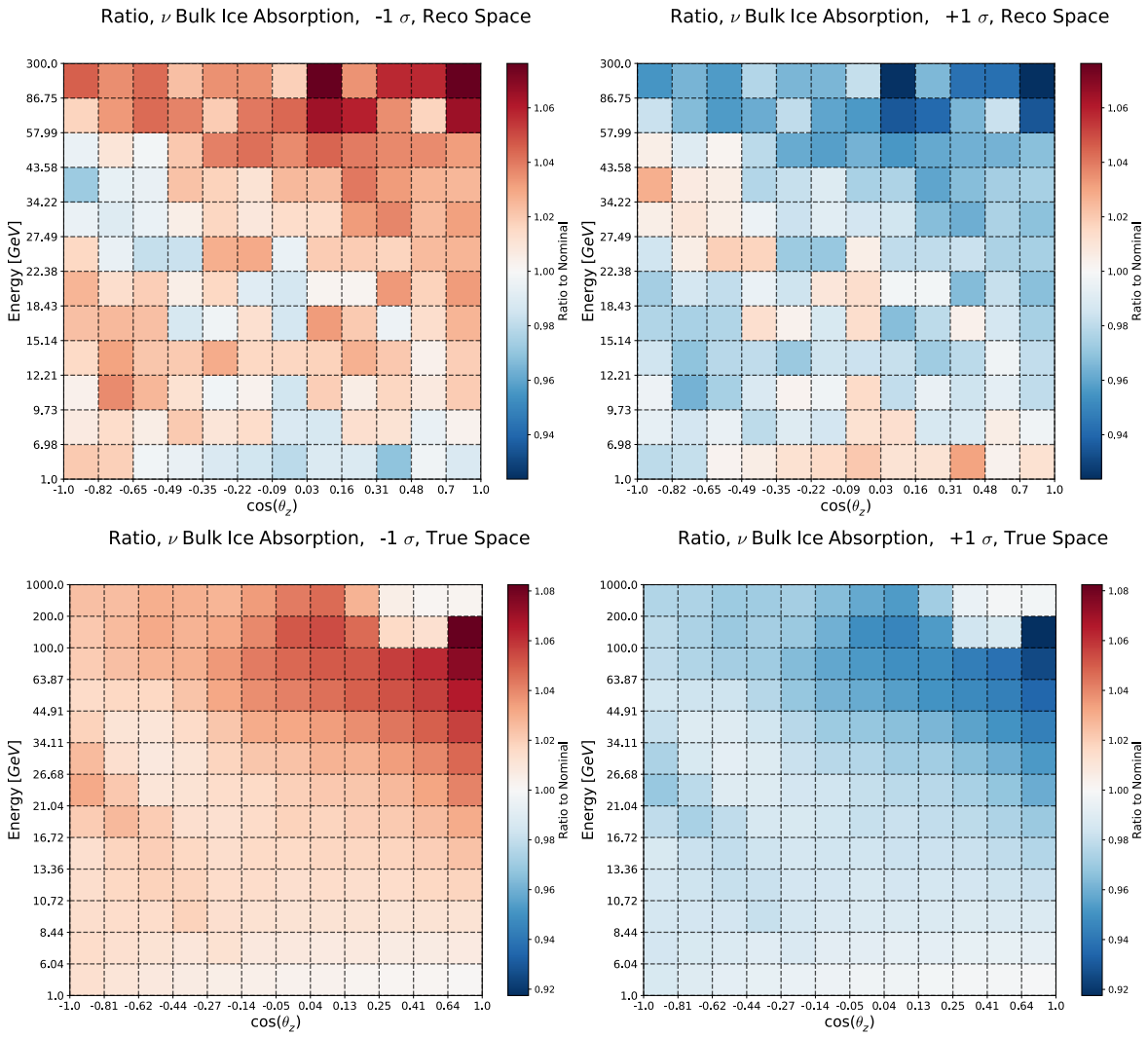


Figure 7.9: Effect of bulk ice absorption on the final level event rates. Upper: reconstructed space. Lower: true space.

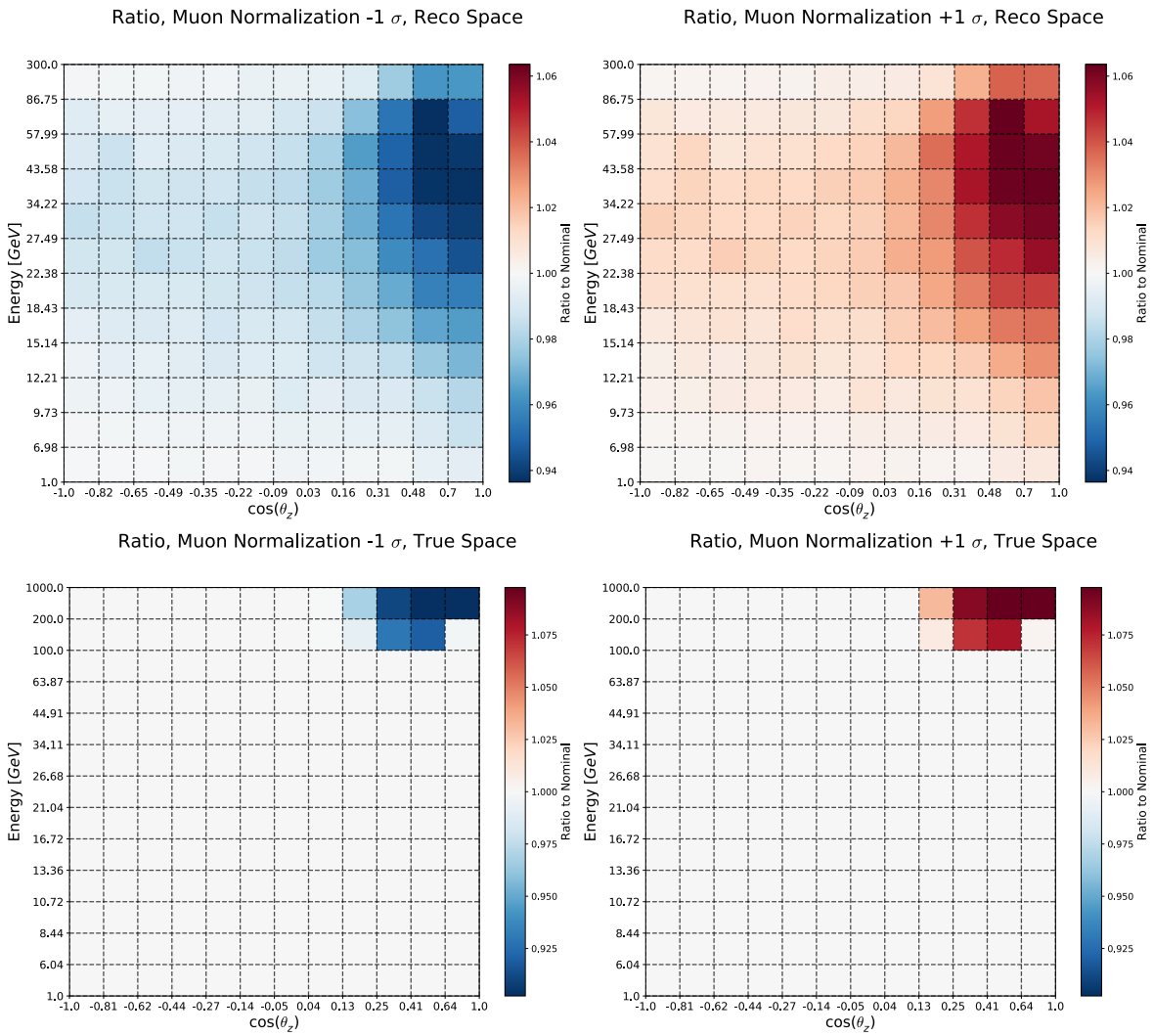


Figure 7.10: Atmospheric muon scale, relative systematic impact at $\pm 1\sigma$ levels. Upper: reconstructed space. Lower: true space.

7.4.3 Noise Scale

A scaling factor is introduced to adjust the level of background noise included in the MC. The noise model is described in sec. 4.4, and the scaling factor simply raises or lowers the overall noise rate. The systematic prior has been set to ±10%, and the impact on the MC event rates are illustrated in fig. 7.11, and lie below the percent level for most of the parameter space.

7.5 Systematic Uncertainty Impact on the Detector Efficiency

In the unfolding scheme we use our MC to generate the response matrix. As we vary a systematic uncertainty in the MC, this will in turn change the efficiency calculation and also the response matrix. This effectively implements the impact from each systematic twice, but since this is an inherent feature in the analysis framework it is impossible to avoid when using this analysis approach under these circumstances. It is thus necessary to know the value of the efficiency under changes to the systematics. This change in the efficiency has been implemented in the analysis software and can be queried in the same way as the impact on the event rate. For illustrative purposes, included here is the efficiency change due to to DOM efficiency systematic being set to its $\pm 1\sigma$ values, as seen in fig. 7.12. The impact is on the scale of approximately 5% to 10% with the largest effects seen for the high-energy downgoing region. Interestingly, a part of the parameter space in the straight upgoing region around $14GeV$ is also sensitive to changes in the DOM efficiency.

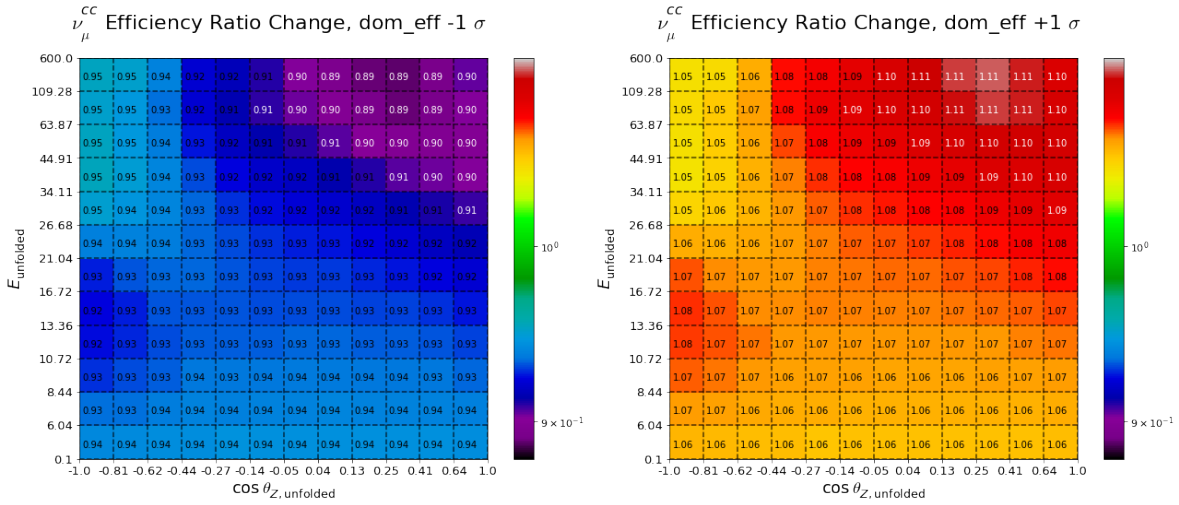


Figure 7.12: Systematics impact on the for the $\nu_\mu^{cc} + \bar{\nu}_\mu^{cc}$ channel, when setting the DOM efficiency parameter to $+1\sigma$ or -1σ respectively

7.6 Systematics Implementation in Analysis

In this work two different approaches have been employed, a bracketing approach and a Gaussian sampling approach. The bracketing approach has merit in the case where several systematic parameters impact the analysis and more than a few have a large effect. This method uses the unfolding performed in three rounds: First for the nominal case, then twice for the cases in which all systematics have

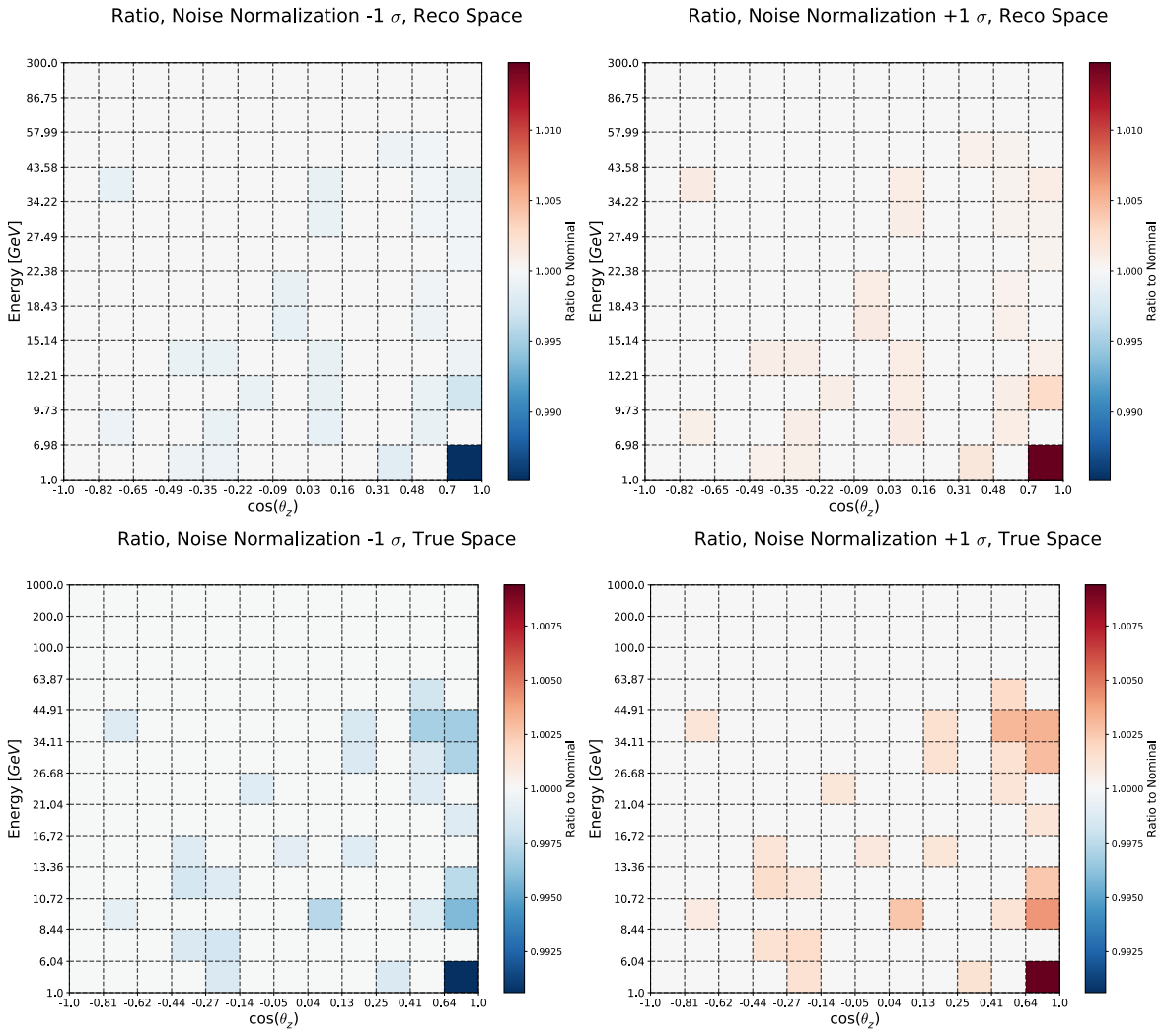


Figure 7.11: Noise normalization, relative systematic impact at $\pm 1\sigma$ levels. Upper: reconstructed space. Lower: true space.

been set to their $\pm 1\sigma$ pull values, as given by their priors. In this approach correlations between the individual parameter's effect on the unfolded event rate has to be taken into account. Most parameters are directly correlated in that increasing their values also leads to a higher event rate, however, some are inversely correlated, where increasing the systematic parameter in question leads to a lower event rate. The systematic uncertainty on the unfolded spectrum is set as the limit of the two $\pm 1\sigma$ extremes. This method can prove to be overly conservative, as the a-priori likelihood for all systematic parameters to have true values at their 1σ priors is unlikely. Furthermore, the bracketing approach explores the 'box' in parameter space bounded by the priors - under the assumption of a Gaussian prior this will be an overestimate, as the multi dimensional shape of a Gaussian prior is elliptical as opposed to the rectangular shape given by a flat prior.

8 Analysis and Results

The unfolding analysis is discussed in this chapter, including: the analysis chain, the closure tests, burn sample tests, and corrections to the binning and final level cuts. We then discuss the application of the analysis method to the unblinded data sample. The chapter begins by exploring the implementation of the Bayesian unfolding method.

8.1 Choice of unfolding dimensions

The neutrinos interacting in IceCube come in three flavours. At best each of these could be resolved perfectly, but due to detector limitations in relation to the light output and kinematics of the events, certain compromises must be made. The incident neutrinos have at the time of interaction, a certain energy, called the *true energy*. Via the interaction this energy is deposited into daughter products and emitted photons are eventually detected by the IceCube DOMs. As described in section 3.5 this process complicates the perfect identification of incoming particles. The *reconstructed energy* is then the energy as calculated by IceCube based on the collect charge on the DOMs. One goal of the unfolding method is compensating for the difference between true energy and reconstructed energy. The same principle applies to the incident angle of the incoming neutrino, resulting in a multi dimensional unfolding problem.

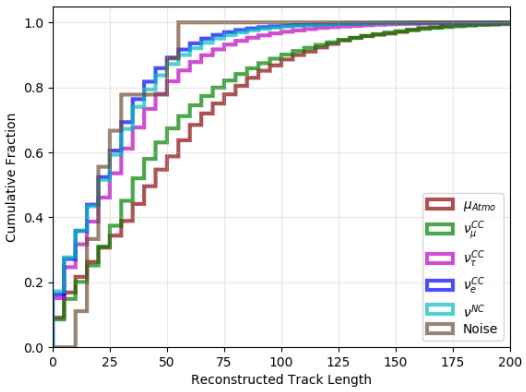


Figure 8.1: Cumulative distributions of the reconstructed track length between interaction types. $\nu_e, \nu_\tau, \nu_\mu^{\text{nc}}$ and noise can be classified as 'cascade-like' while ν_μ^{cc} and atmospheric Muons can be classified as 'track-like' [60]

When performing an unfolding in multiple dimensions, it is important to realize what those dimensions are. Here, we will discuss the decisions regarding the choice of unfolded channels. The goal of the unfolding is to compensate for detector effects and relate the reconstructed event information to the underlying true values of nature. Four of the analysis dimensions are: reconstructed energy, reconstructed zenith angle, true energy, and true zenith angle. As discussed in section 3.5, events are expected to leave a track, a cascade, or a double bang signature in the detector, depending on the flavor of the incoming neutrino, the energy, and the interaction type. However, due to detector design, particle identification in IceCube can prove difficult, as GeV-scale Muon Neutrino events will leave only a short muon track, and thus display characteristics similar to those of neutral current cascades encompassing all flavors. The unfolding procedure can alleviate this issue. It is possible to also unfold by flavour and interaction type, and correlate observed event topology with true flavor and interaction. This is done by introducing a measure of 'trackness', and by classifying events as 'track-like' or 'cascade-like', based on the reconstructed track length. Using the GRECO sample the Pegleg reconstruction yields a reconstructed track length L_r, allowing for the following categorization:

$$\text{Track} - \text{like} : L_{\text{r}} \geq 50m \tag{8.1}$$
$$\text{Cascade} - \text{like} : L_r < 50m \tag{8.2}$$

This characterization is motivated by the cumulative distribution of reconstructed track length amongst the different flavors as shown in fig. 8.1. Under the assumption that the outgoing muons in charged current interactions are minimally ionizing, ideally the expected track length is 4.5m per GeV: $L = E\left(\frac{1.0GeV}{4.5m}\right)^{-1}$, resulting in ~50m for a 10 GeV particle. The 50m separation has thus been chosen by the maximum difference between the cumulative distributions of true 'track' events and true 'cascade' events, as seen in fig. 8.1. For the corresponding truth dimension, we select the only channels expected to leave a clear track signature in the detector: $\nu_\mu^{\text{cc}} + \bar{\nu}_\mu^{\text{cc}}$ for one channel and all other flavour and interaction types for another channel. The pros and cons regarding channel choice have been carefully considered, and a short summary will be given here: In principle it would be possible to unfold the different flavours individually and thus group together all muon-, electron and tau interactions into three unfolding bins. Firstly, this would lead to a shortage of event statistics, particularly in the ν_τ bin, as our MC is limited at present to 491094 events (see section 6.6 for a discussion of MC sample size requirements). Secondly, this would constitute a measurement of the ν_τ normalization, which would depend strongly on the oscillation parameters. Thirdly the unfolding method is imperfect and the small absolute difference between events in the ν_e and ν_τ channels would leave the measurement strongly dependent on systematic errors, eg. the DOM efficiency. By keeping the $\nu_e, \bar{\nu_e}$, $\nu_\tau, \overset{(-)}{\nu_\tau}$ and $\nu_\mu^{\text{nc}} + \bar{\nu}_\mu^{\text{nc}}$ signals in the same unfolding bin, the direct dependence of the event rate on the oscillation is thus reduced. This leaves the unfolded truth side of the unfolding matrix with a true 'track' channel consisting of $\nu_\mu^{\text{cc}} + \bar{\nu}_\mu^{\text{cc}}$ events and one true 'cascade' channel consisting of all other flavor and interaction types.

8.2 Number of Iterations

The number of iterations in the unfolding represents how many times the output from the unfolding is reapplied to update the Bayesian prior and unfold again. It is the Bayesian unfolding pendent to the regularization parameter occurring in other unfolding schemes, and its main function is to overcome the bias arising from the initial prior. The unfolding framework is built from our nominal Monte Carlo simulation, so in case of few iterations it will be biased towards this initial prior, while at higher iterations the unfolded spectrum will converge on the true value. However, with more iterations fluctuations in the measured spectrum can grow and be blown out of proportion - this applies both to statistical and systematic deviations. Generally this type of unfolding converges quickly, giving a rapid improvement in the maximum likelihood estimate, within 3 to 5 iterations [72, 70]. In this section we will discuss how to balance the bias vs. the uncertainty to set the optimum number of iterations.

8.2.1 Uncertainties and Iterations

As the unfolding method is iterative and takes the output from one step as input to the next, there is a very real risk of uncertainties becoming inflated, as they can grow every time a statistically uncertain quantity is used as input. This effect is actually two-fold as statistical fluctuations can also grow with the number of iterations. It is thus necessary to limit the number of iterations at a small enough value that these fluctuations and systematic effects are still small. However, there is the caveat that if too

few iterations are chosen, the unfolded sample will be biased towards the Monte Carlo simulation used to construct the unfolding framework. The growth of statistical and systematic effects with number of unfolding steps must be evaluated. A preliminary investigation of this effect was implemented by a systematic shift in the input spectrum. Section 8.3 will be a small intermezzo to discuss this.

8.3 Pseudo Measurement by Modifying the spectral index

A spectral index shift, $\delta\gamma$, is included as a modification to the flux based on the energy, due to the spectral index power law dependence. Several values of $\delta\gamma$ are tested by random sampling, resulting in pseudo experiments, which are systematically and statistically different from the nominal Monte Carlo. Two versions of the spectral index modification were implemented and tested, and both will be discussed in the following.

8.3.1 Standard Spectral Index Shift

The nominal spectral index is shifted by a small amount $\delta\gamma$, thus we expect the endpoint of the flux distribution as well as the normalization of the flux to vary. The standard modification is defined as follows:

$$\Phi'_{\nu_x} = E^{\delta\gamma}\Phi_{\nu_x}, \tag{8.3}$$

when just modifying the spectral index directly as a function of energy we thus expect a modification of the final event rate we see in our pseudo experiment. The standard systematic prior on the spectral index is $\delta\gamma = \pm 0.05$, and several samples have been produced at values up to three times this standard value. as illustrated in fig. 8.2.

8.3.2 Spectral Index Shift - With tipping point

The spectral index shift, $\delta\gamma$ is included as a modification to the flux based on the energy, due to the spectral index power law dependence. Considering our binning running in irregular bins from 0.1GeV to 600GeV a tipping point was introduced at 27GeV and the modified flux was defined in the following way:

$$\Phi'_{\nu_x} = A\left(\frac{E}{27\text{GeV}}\right)^{\delta\gamma}\Phi_{\nu_x}, \tag{8.4}$$

where A is an integral preserving constant and $\delta\gamma$ is the shift in the spectral index. When the energy is above the tipping point at 27 GeV, the fraction raised to the power of $\delta\gamma$ will be larger than one, giving an increase in the flux, and vice versa when the energy is below 27 GeV. The value $27GeV$ was chosen as it is roughly in the center of the logarithmic binning range. We thus expect the overall event rate to be constant, but events to be redistributed around the tipping point, as can be seen in fig. 8.3.

8.3.3 Systematic Prior on the Spectral Index

Knowledge of the spectral index comes mainly from measurements of primary cosmic rays - specifically results from AMS-02 were used by Honda et al. [9] to calibrate their neutrino flux model at higher energies, where the index defines the endpoint. Looking at data from various sources, plotting the Cosmic Ray flux used as input to the HKKM simulation as well as a shift in the spectral index,

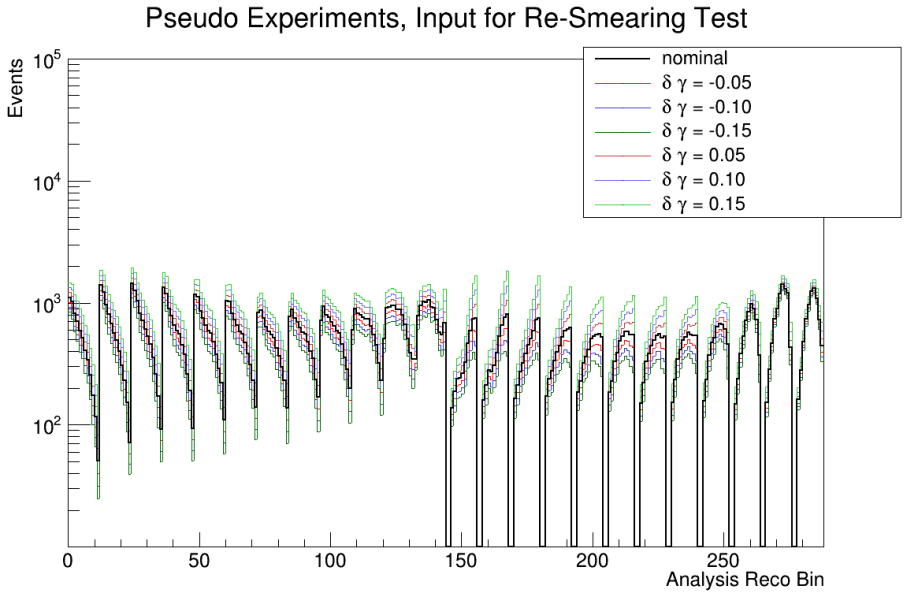


Figure 8.2: Pseudo Measurements obtained by changing the spectral index according to eq. 8.3. Shown is the flattened spectrum in zenith angle, so that every 12 bins constitute the energy spectrum for one zenith angle direction. Each pseudo measurement corresponds to a specific change in spectral index. As expected we see a change in event rate, and only a small change in the shape of the spectrum.

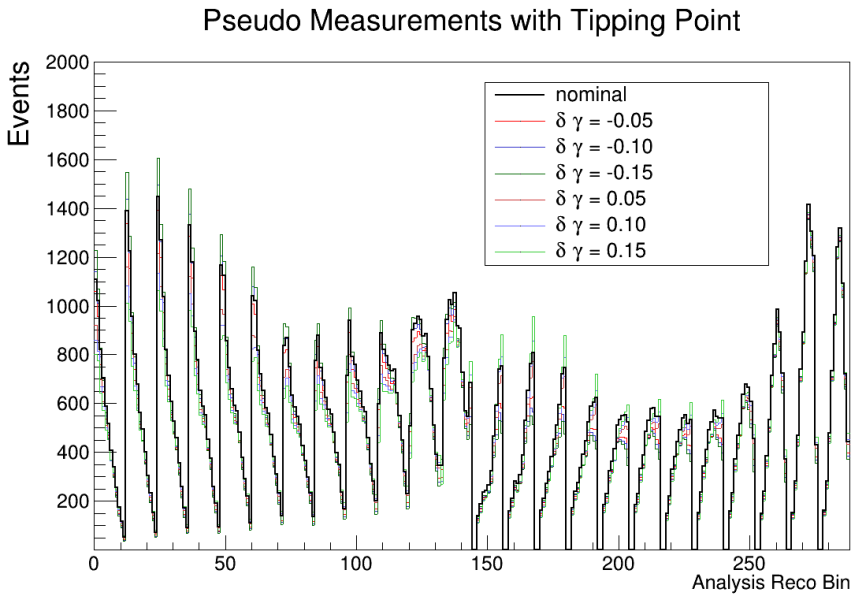


Figure 8.3: Pseudo measurements used as input for the re-smearing test. Shown is the flattened spectrum in senith angle, so that every 12 bins constitute the energy spectrum for one zenith angle direction. The spectra shown are for various shifts in the spectral index, using a tipping point at $27GeV$, and a rate preserving constant as per eq. 8.4. As expected the shape of the spectrum is changed, while the overall event rate is constant.

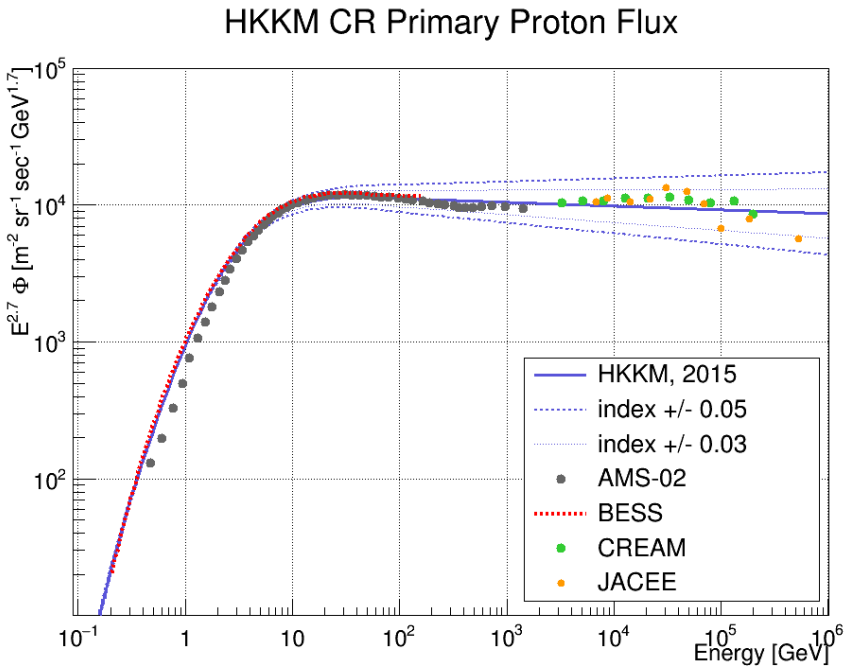


Figure 8.4: Primary Cosmic ray flux used as input for the HKKM flux model, as well as recent data from various experiments. The HKKM-2015 model uses a spline to the BESS data up to ~100 GeV after which a power law with index $\gamma = -2.68$ is used [9]. The HKKM authors are currently working on a new version of the HKKM atmospheric neutrino flux simulation which uses the latest measurements from AMS-02 and is calibrated by muon measurements. This is as of yet unpublished and we use the HKKM-2015 model in this work.

corresponding to the standard systematic prior of $\delta\gamma = \pm 0.05$, as illustrated in fig. 8.4, shows this prior to be overly conservative in our energy range of interest (0.1GeV to 600GeV). The only issue comes below a few GeV where a straight modification of the flux without a tipping point does not span the difference between measurements. The HKKM model authors have been working on an updated version of the neutrino flux model, using a spline to the AMS-02 data below 100 GeV, and a power law with index $\gamma = -1.68$ above 100 GeV, as well as muon data to calibrate the normalization [9]. Unfortunately this has not yet been published, therefore this analysis makes use of the HKKM, 2015 model [7]. In general the systematic prior on the spectral index could reasonably be tightened to a value of $\delta\gamma = \pm 0.03$. With this in mind the shifts $\delta\gamma$ out to $\pm 3\sigma$, used above, extremely conservative. It is worth noting however, that the models are slightly different due to the tipping point and integral preserving factor used above. Even so, checking the $\pm 3\sigma$ range is more than adequate for investigating the impact on an unfolding analysis of a shift in the spectral index.

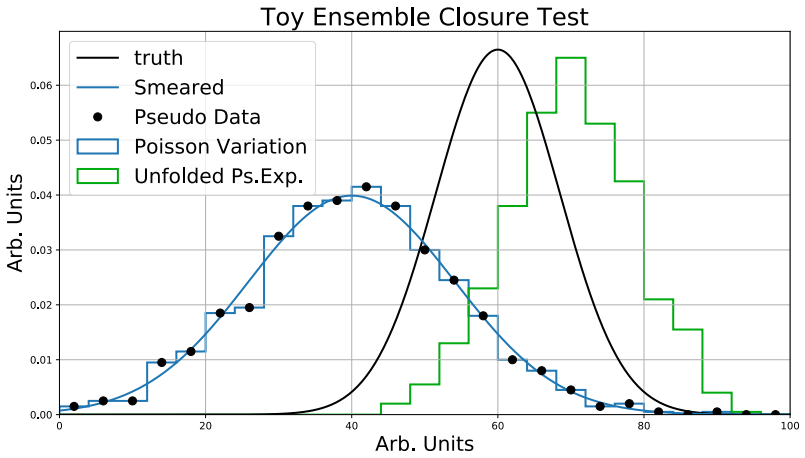


Figure 8.5: Illustrating the unfolding principle with a pseudo measurement input. The true MC distribution is shown in black, with the response-smeared distribution in blue. From the blue distribution 200 pseudo experiments are drawn, via a Poisson variation around the bin value. One such is shown in the blue histogram, with the black dots used as pseudo data input for the unfolding. The resulting unfolded histogram is shown in green.

8.4 Unfolding Ensemble Test

In order to test the impact of statistical and systematic uncertainties and their evolution with the number of iterations, two instances of an ensemble test are performed, using the same procedure as described in section 5.4, but in a different software implementation. Two hundred pseudo experiments are drawn around the nominal reconstructed Monte Carlo simulation, as well as around a reconstructed spectrum with a shifted spectral index as discussed above in section 8.3. Each pseudo experiment is then unfolded using the standard unfolding method with all systematics set to their nominal values. An illustration of the spectra used in this ensemble test is shown in fig. 8.5 The unfolded spectra are then compared to the MC-truth via a simple χ^2 test. The reduced χ^2_{red} for each iteration is defined from the unfolded bin value u_i, the true bin value t_i and the uncertainty in bin i σ_i as:

$$\chi^2_{\text{red}} = \frac{1}{N} \sum_{1}^{N} \left(\frac{u_i - t_i}{\sigma_i} \right)^2 \tag{8.5}$$

and is plotted as a function of iteration steps, as shown in fig. 8.6, yield results similar to the preliminary test performed in section 5.4, in spite of the software difference. As expected the nominal case converges in the first iteration after which it cannot improve the match to the truth. The shifted spectrum on the other hand sees a large deviation from the truth after the first iteration, but undergoes a rapid increase in consistency between unfolded and truth with the first few iteration steps. A plateau is reached in the χ^2 distribution after about 15 iterations. These tests confirm the expected behavior of the unfolding. The challenge from here on then becomes the optimization of the number of iterations.

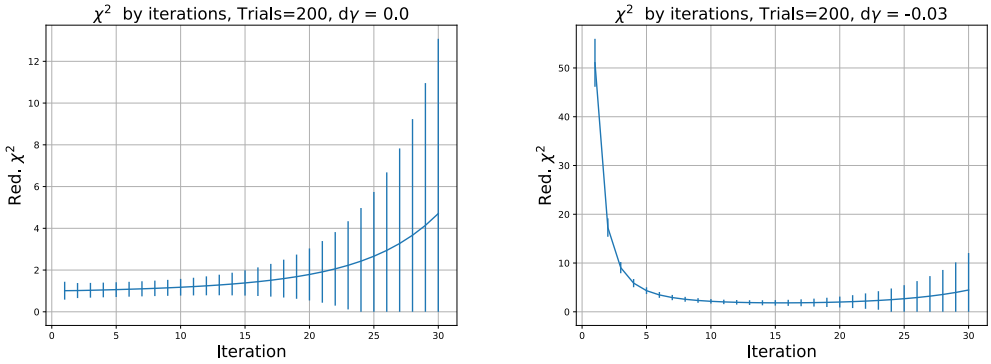


Figure 8.6: Average reduced χ^2 as a function of iteration steps, for the two preliminary ensemble tests. The error bars correspond to the variance among the 200 χ^2s. Left panel shows the nominal case where the unfolding returns the best match to the truth in the first iteration and iterating mostly makes statistical fluctuations grow. Right panel shows the case where the reconstructed spectrum is shifted by $\delta\gamma = -0.03$, and the unfolding shows a rapid increase in consistency with the first few iterations. A plateau in the χ^2distribution is reached after ~10 iteration steps.

8.5 Evolution of the Statistical Uncertainty

The statistical fluctuations in our sample can be enhanced by the unfolding, and the statistical uncertainty will grow as it is fed into the unfolding repeatedly with each iteration step. In order to check the impact of this the unfolded spectra from the ensemble test above in sec. 8.4 are averaged for each step of the unfolding. As an example the event rate for a single energy bin is plotted as a function of iterations in fig. 8.7. The unfolding converges as the unfolded event rate is constant with the number of iterations, however, the size of the error bars increases with the number of iterations. The behavior is plotted in the right panel in fig. 8.7, where the different lines illustrate the evolution of the absolute statistical uncertainty in various energy bins. Even though the error bars look large and the evolution can seem drastic, the effect is modest - reaching only about 5% relative uncertainty in the worst case scenario of a low statistics bin and 24 iterations.

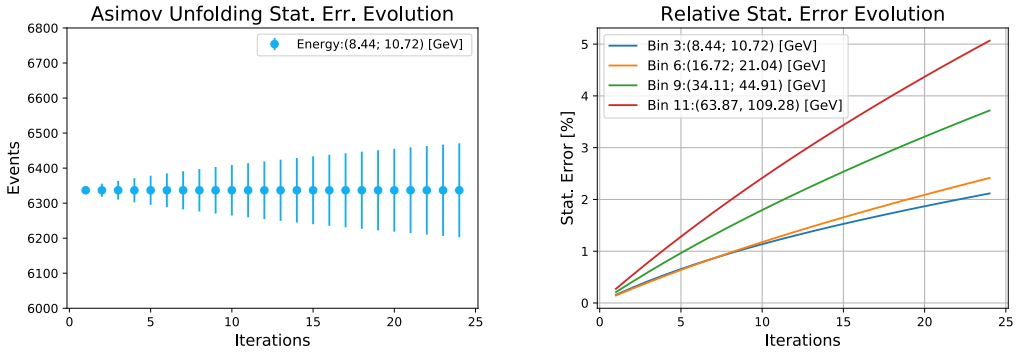


Figure 8.7: Evolution of the statistical uncertainty in the unfolding. Left: Unfolded event rate as function of iterations, averaged over 200 pseudo experiments. The distribution is flat indicating the unfolding converges as expected at the first iteration. The error bars are seen to grow with the number of iterations, clearly illustrating the enhancement effect of the iterative unfolding procedure on the statistical uncertainty. Right: The evolution of the statistical uncertainty in number of events, with the number of iterations for various energy bins.

8.5.1 Statistical Uncertainty Method Consistency Check

In order to verify the above propagation of the uncertainty, a check was implemented to compare the statistical uncertainty evolution as a function of number of iterations, across implemented methods. It is important to realize the statistical uncertainty error bars represent the range within which 68% of all unfolded random trials should fall. It is thus possible to verify the matrix-propagation method described in section 6.5 by comparison to uncertainties calculated from an ensemble of unfolded pseudo experiments. RooUnfold has a built-in method for testing this, using an ensemble of 20 pseudo experiments [72]. The calculation of the uncertainty is performed as the square root of the covariance matrix of the unfolded pseudo experiments. To further verify the evolution of the statistical uncertainty, an independent method was implemented as well, in which the statistical uncertainty is simply taken as the square root of the variance of the unfolded values. The average statistical uncertainty over all bins as a function of iterations is shown in fig. 8.8 and shows good consistency between all three methods. As expected the uncertainty grows with the number of

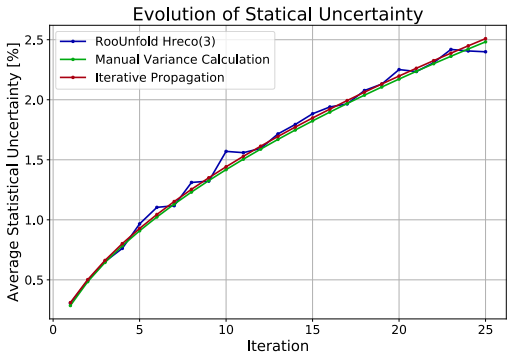


Figure 8.8: Evolution of average relative statistical uncertainty with number of iterations for three different calculation schemes. RooUnfold Hreco(3), shown in blue, is the built-in method in RooUnfold to estimate statistical uncertainties from an unfolded ensemble. A manual calculation using a 200 trial ensemble is shown in green, and finally the matrix-propagation method is shown in red. Good consistency is seen between all three methods.

iterations, and while the starting value is below 0.5% the iterative procedure can push this up to 2.5% within the 25 iterations tested. The RooUnfold native method is displaying a less smooth curve due to the relatively low number of pseudo experiments, namely 20. The comparisons performed in this test confirm the fast matrix-propagation method to be consistent with expectations and this method will be employed henceforth.

8.6 Choice of Test Statistic and Stopping Condition

In order to evaluate the consistency between the unfolded spectrum and the truth, a measure of similarity needs to be introduced. So far we have made use of the simple Pearson's χ^2 definition:

$$\chi^2 = \sum_i \left(\frac{u_i - t_i}{\sigma_i} \right)^2, \tag{8.6}$$

with u_j and t_j being the bin content of the unfolded and true histograms respectively. σ here represents the statistical uncertainty on the unfolded spectrum. We expect values of this χ^2 to be close to one, insofar as the unfolded histogram bin values represent a random sampling from a Poisson distribution. This χ^2 definition, while measuring the bin-by-bin difference between truth and unfolding in terms of the statistical uncertainty, gives no indication of whether that uncertainty is large or small in absolute terms. Under the condition that $u_j - t_j$ is roughly equal to σ, the absolute scale of σ will not be represented in the numerical value of χ^2. This is a weakness of this χ^2 definition. As the statistical uncertainty grows with the number of iterations, while the difference between unfolded and truth becomes smaller, the test statistic is expected to rapidly drop with the number of iterations. At high iterations where the unfolding can diverge, the test statistic is expected to grow, insofar as the shift becomes larger than the statistical uncertainty. Both of these effects are observed in sec. 8.4 and illustrated in fig. 8.6.

As the main analysis presented in this thesis progresses from an Asimov test, to MC ensemble tests and onward to burn sample tests including systematics, it is prudent to alter the choice of test statistic. A choice of test statistic that encodes the similarity of the two distributions is the χ^2 definition built-in to ROOT:

$$\chi^2 = \frac{1}{UT} \sum_i \frac{(Ut_i - u_i T)^2}{u_i + t_i}, \tag{8.7}$$

where U and T are the total number of events in the unfolded and true histograms respectively, while u_i and t_i represent the number of events in bin i in the unfolded and true histograms respectively. This test statistic definition does not include a term for the statistical uncertainty. As such it contains different information than the definition in eq. 8.6, as it instead includes the integrated spectrum. Because the denominator now depends on the event rate instead of the statistical uncertainty, this measure is not prone to reduction of the test statistic due to growth of statistical uncertainties. This definition is used in many of the ensemble tests which will be discussed in the rest of this chapter.

Setting the stopping condition to the optimal value is a trade-off between bias in the unfolding on the one hand, and uncertainties on the other hand. For the Asimov unfolding tests, very little to no bias is observed, and the above definition of the test statistic in eq. 8.7 is used. In these cases it makes sense to truncate the unfolding when the difference from one iteration to the next is no longer statistically significant: when the difference in test statistic from the i'th iteration to the i-1'th iteration

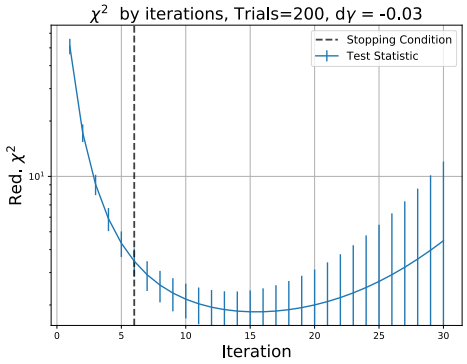


Figure 8.9: Test statistic as function of iteration. This plot corresponds to the right panel of fig. 8.6, but with the y-axis in log scale to better illustrate the behaviour of the reduced χ^2 as related to the size of the error bars with the number of evolutions. The vertical black dashed line indicates the stopping condition from eq. 8.8.

drops below the width of the TS error bar σ_i, this condition is fulfilled. Thus that stopping condition can be expressed as:

$$\Delta \mathrm{TS_i} \leq \sigma_i, \tag{8.8}$$

with i indexing the iterations, σ_i being the statistical uncertainty and $\Delta\mathrm{TS} = \mathrm{TS_i} - \mathrm{TS_{i-1}}$. The stopping condition for the unfolding ensemble test discussed in section 8.4, is shown in fig. 8.9. This definition of stopping condition does not inherently take into account the statistical uncertainty in the individual bin, however, as the binning has been optimized to keep the statistical uncertainty around the few percent level, this is still reasonable - indeed the consistency checks in section 8.5.1 reveal the relative statistical uncertainty to be below 2.5%, even at 25 iterations.

The TS calculation takes into account the difference between unfolded and simulation truth, but only does so in absolute terms. The goal here is to make the unfolding as precise as possible - the iterative procedure should be truncated before the iterations bring the unfolding into a regime with either a large difference between unfolded and truth σ_{unf}, large statistical uncertainties σ_{stat} or both. In this regard finding the optimum stopping condition again becomes a trade-off between unfolding pull on the one hand, versus uncertainties on the other. It is not justifiable to keep iterating once the unfolding pull becomes smaller than the error in the measurement. These considerations then lead to setting the TS to be the envelope of the uncertainty band on the measurement:

$$\mathrm{TS}_\sigma = \frac{1}{N}\sum_i \left(\sigma_i^{up} - \sigma_i^{dn}\right)^2, \tag{8.9}$$

where $\sigma_i^{\mathrm{up}}$and σ_i^{dn} are the upper and lower edges of the 68% confidence interval of the measurement, with σ_i defined as follows:

$$\sigma_i^2 = \sigma_{\mathrm{stat}}^2 + \sigma_{\mathrm{sys}}^2 + \sigma_{\mathrm{unf}}^2 \tag{8.10}$$

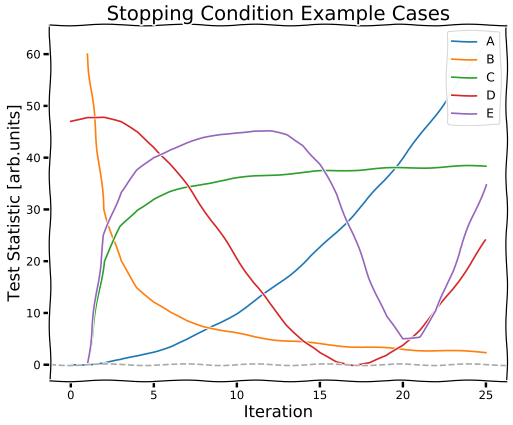


Figure 8.10: Hypothetical evolutions of the test statistic with the number of iterations.

The stopping condition is then given as the iteration corresponding to the minimum of the envelope TS:

$$n_{\text{stop}} = i\left(\min\left[TS_{\sigma}\right]\right) \tag{8.11}$$

With the test statistic defined, a more general stopping condition must be imposed. It is generally favorable to minimize the test statistic, except in cases where the evolution of the test statistic is divergent. Figure 8.10 discusses the hypothetical scenarios of the TS evolution with the number of iterations. Depending on the shape of the distribution the strategy for setting the stopping condition will be slightly different. A short discussion of the different cases follows. In case A, the stopping condition value sharply increases and does not converge, showing a minimum at one iteration (fig. 8.10). As using just one iteration would introduce a bias in the unfolding towards the MC this should be avoided. The initial increase will be disregarded and the number of iterations will then be set to five, as prescribed by D'Agostini [70]. In case B, the distribution converges towards zero and the number of iterations will be set to the value corresponding to the minimum of the curve within the sampled number of iterations, in this case 25. In case C the distribution sharply increases before tending towards a plateau. For similar reasons as in case A the initial increase will be disregarded and the stopping condition will be set at the point where the derivative of the curve drops below 1% of the TS value. In case D where the distribution displays several minima, the stopping condition will be set to the minimum corresponding to the lowest TS value. In case E, the initial increase will be disregarded for the same reasons as in case A and C. The distribution at iterations above the increase will then be minimized according to the logic from case D. The cases are summarized in table 8.1.

8.6.1 Statistics Only Unfolding

Based on the ensemble tests performed in section 8.4, it is now prudent to explore what an actual unfolded result would look like. To accomplish this, a preliminary closure test is performed without weighting the prior term. A pseudo experiment is drawn from the standard MC via a Poisson variation around the individual bin values, and unfolded using the D'Agostini Bayesian unfolding method

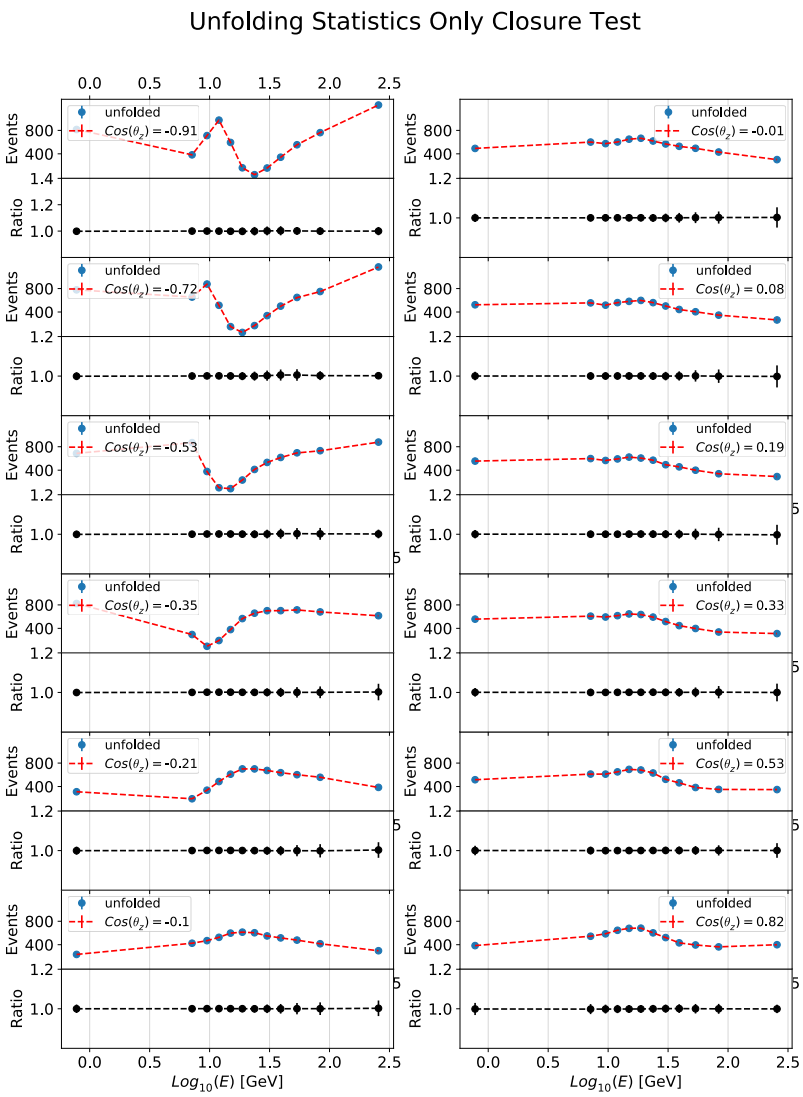


Figure 8.11: Statistics only closure test results for each $\cos(\theta_z)$ bin. The error bars are small and represent statistical uncertainty only. The oscillations impact can be clearly seen for the upgoing bins. Good agreement is seen for all bins to within the few percent level.

Case	Logic
A	Stop after 5 iterations
B	Global Minimum
C	Disregard initial increase, set by 1% change.
D	Global Minimum
E	Disregard initial increase + Global Minimum

Table 8.1: Logic of setting the stopping condition for the hypothetical cases shown in fig. 8.10.

described in chap. (5.2). Five iterations are used and the unfolded results for each $\cos(\theta_z)$ bin are shown in fig. (8.11). All unfolded bins show good agreement with the input MC and the statistical uncertainty is in the few percent range for all bins. This gives a strong indication that the analysis is systematics dominated. With this test done, a preliminary comparison of the statistics only unfolding results to the Super-K atmospheric flux unfolding can be performed. This is covered in the next section.

8.6.2 Expectation: Comparison to Super-K

Using the standard binning, a pseudo experiment is drawn from the reconstructed MC and unfolded using the standard 5 iterations from D'Agostini [70]. Including no systematic uncertainties, it is possible to get an early expectation of the validity of this measurement as compared to Super Kamiokande results [3]. For the two PID channels, the unfolded spectra relative to the input in the statistics-only case is shown in fig. 8.12. As expected the statistical uncertainty on the unfolded values is on the percent level and measurements IceCube should be highly competitive in this field, provided the systematic uncertainties can be sufficiently controlled.

8.7 Re-Smearing Closure Test

As mentioned above, in the case of few iterations (e.g. 1-2), the unfolded result might be biased towards the initial prior. However, when testing for this, we must factor in that we do not a-priori know the specific properties of the measurement to come. We will just measure some event spectrum and proceed to unfold it. We can simulate this scenario by generating pseudo experiments which are statistically different from the standard Monte Carlo and then use those as pseudo measurements to which the unfolding method is applied. This will allow a closure test of the unfolding with regards to the compatibility of the unfolded spectrum, even when the truth which gave rise to the measurement is unknown (indeed - this is what we are trying to estimate!). The procedure goes as follows:

1. Generate an input pseudo measurement - based on the reconstructed MC sample. (Blue in fig. 8.13)
2. Draw N pseudo experiments around the pseudo measurement - in this case we will use 200. (Black dots in fig. 8.13)
3. Unfold each pseudo experiment into true space (Green in fig. 8.13)
4. Smear the unfolded pseudo experiments by the response matrix used - this is then named the re-smeared histogram (Orange in fig. 8.13)

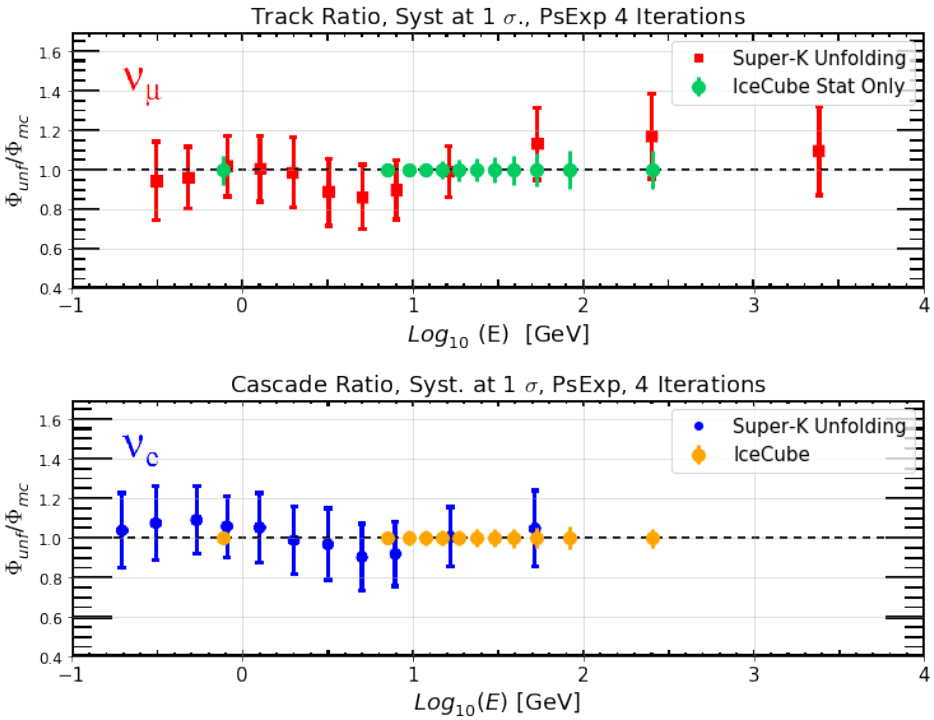


Figure 8.12: Unfolded pseudo experiment shown relative to input MC, marginalized over the $\cos(\theta_z)$ dimension. Upper: ν_μ^{cc}, Lower: $\nu_e + \nu_\tau$. The IceCube data points all fall close to one, meaning the unfolding is consistent. The error bars show statistical uncertainty only, and are low compared to the Super-K unfolding results, indicating this analysis to be systematics dominated.

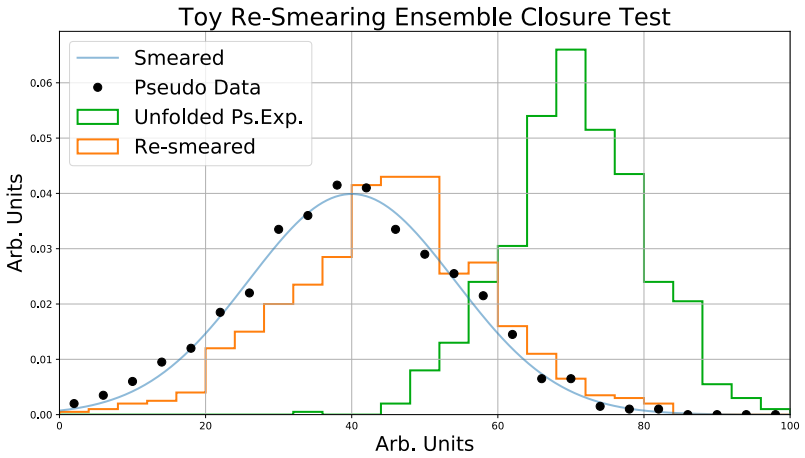


Figure 8.13: Illustrating the principle of the re-smearing test. One of the 200 pseudo-data inputs is shown in black dots and the unfolded histogram is shown in green. Each unfolded histogram is then re-smeared via the response matrix, and is shown here in orange. Each re-smeared histogram is then compared to the smeared distribution from which the pseudo-measurements are drawn.

5. Do a test statistic comparison of the re-smeared histogram to its corresponding pseudo measurement for each unfolding step - in this case up to 25 steps.

6. Average the test statistic from the 200 comparisons between unfolded spectra and pseudo experiments

7. Optimize the number of iterations by minimizing the average test statistic

The process is named re-smearing as a first detector smearing is included in the MC-sample.

8.7.1 Re-Smearing Toy MC Consistency Test

In order to verify the above algorithm and confirm its behavior, a toy MC study is implemented. A response matrix is generated by sampling two different Gaussian distributions to give both a smearing and a systematic shift from true space to reconstructed space. A curve with different parameters is then used as a simulation of the underlying truth of nature. All these distributions are illustrated in fig. 8.14. The true and reco distributions are then randomly sampled 200000 times to generate a binned version of the response matrix as well as true and reco histograms corresponding to the axis projections of the response matrix. These histograms are illustrated in the two top panels of fig. 8.15. The following closure tests are then performed: firstly the response matrix is applied manually to the truth histogram, to verify that this procedure correctly yields a result compatible with the reco histogram of the response matrix. This is illustrated in the red dashed line in the upper left panel of fig. 8.15. The smeared histogram is then unfolded using the standard D'Agostini implementation for five iteration steps. The resulting histogram is illustrated by the green dashed line in the upper right panel of fig. 8.15. As a final closure test a true spectrum is generated by discretizing the truth

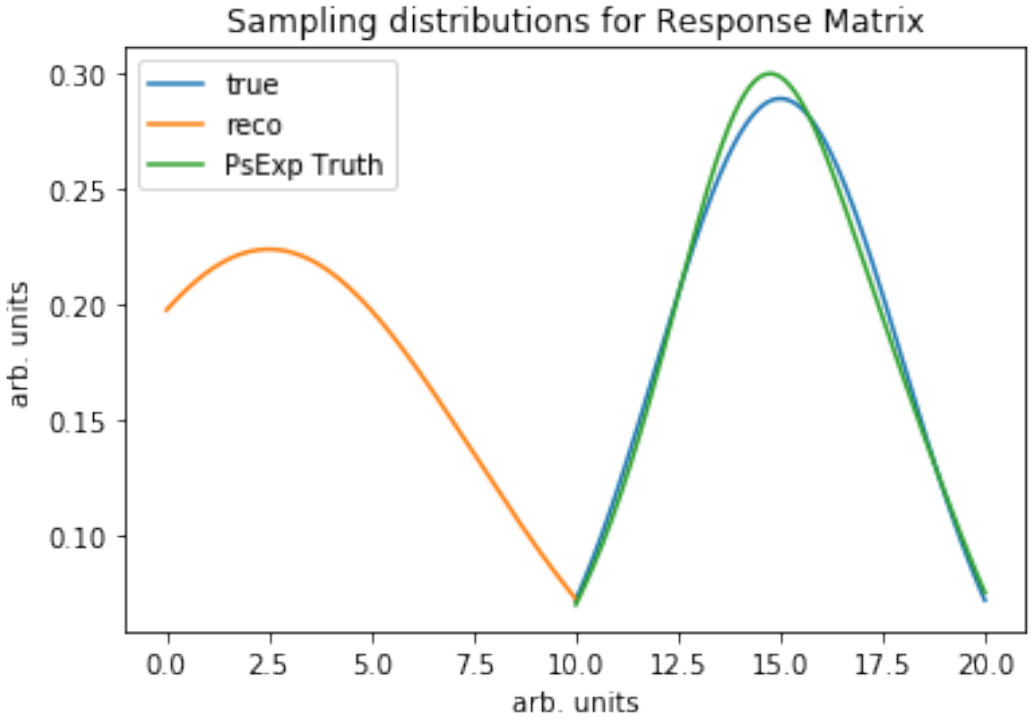


Figure 8.14: Toy MC distributions used to generate detector response and natural truth. Blue: Assumed prior knowledge of natural truth, which goes in the truth side of the response matrix. Orange: Reconstructed spectrum which goes in the reco side of the response matrix. Green: Independent distribution used as natural truth to be 'discovered' by the unfolding. The generated response matrix thus describes the detector smearing from truth (blue) to smeared (orange).

distribution in fig. 8.14 and weighting it to 391 events. It is then smeared and unfolded using five iteration steps. The discretized and unfolded histograms are shown in the bottom right panel of fig. 8.15. We see good agreement in all closure tests, this implies the toy MC is implemented correctly and the software behaves as expected.

8.7.2 Uncertainties and Iterations

Here we investigate the impact from the uncertainties as a function of the number of iterations. In this case we take the nominal Asimov pseudo measurement as input and perform the re-smearing test from sec. 8.7 for each systematic parameter individually. This will reveal which systematics contribute the most to the uncertainty, which we already have good indication of from the investigations in chapter (7). However, this section will give an indication of the degree to which the unfolding can compensate for the systematics impact. The systematic uncertainty on the unfolded measurement as a function of iterations is shown for each systematic parameter in figures 8.168.178.188.19. A convergent behavior is observed in certain bins, while in others a divergence from the pseudo measurement is observed. Most

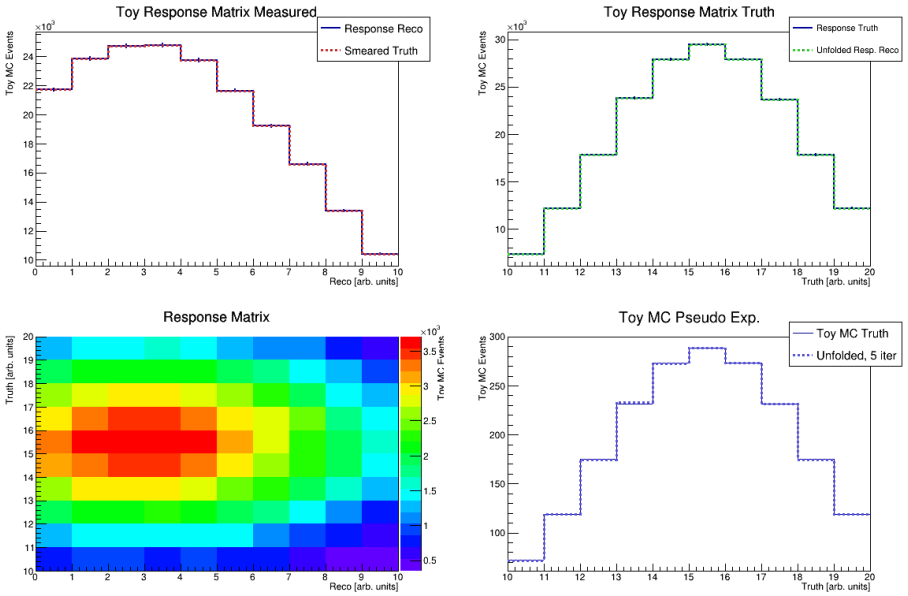


Figure 8.15: Response Matrix and Closure tests of the Toy MC implementation. Top Left: Reco histogram of the response matrix and smeared truth of the response matrix. Top right: Response matrix truth and unfolded histogram of the smeared truth spectrum. Bottom left: Response matrix as used in this toy mc test. Bottom right: Pseudo experiment drawn around the reco spectrum, unfolded with 5 iteration steps. We see good agreement in all cases.

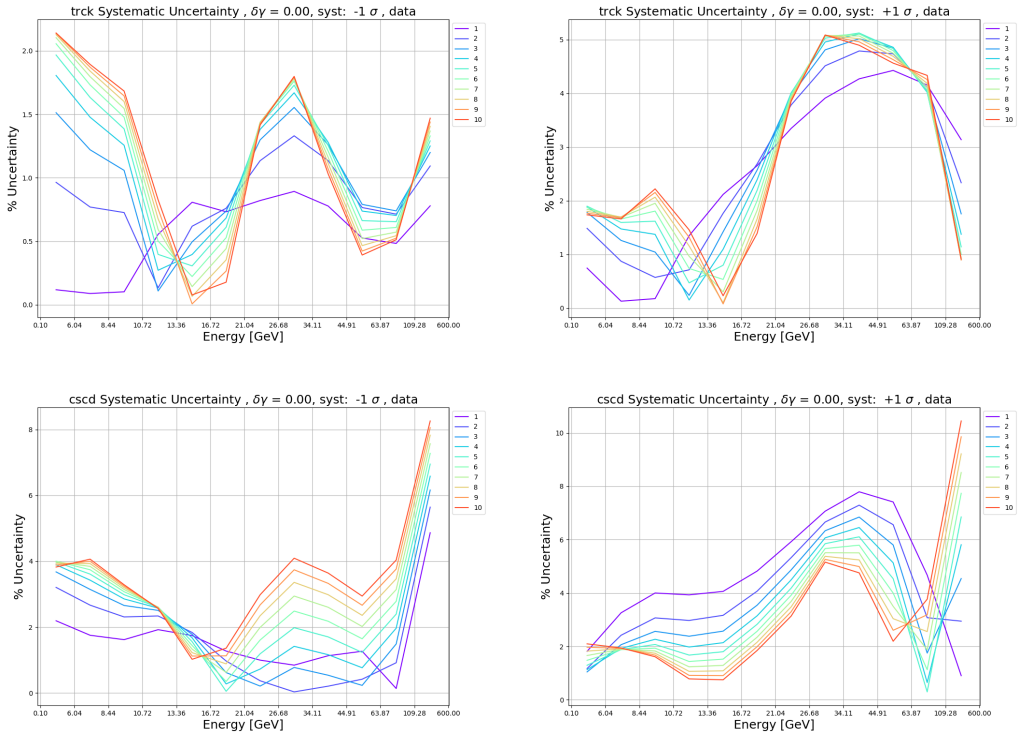


Figure 8.16: The relative impact of the DOM efficiency in the re-smearing unfolding test. Top: track channel. Bottom: Cascade channel. Left: DOM efficiency at -1σ. Right: DOM efficiency at $+1\sigma$

of these systematic uncertainties fall in the $1-10\%$ range, however particularly the cascade channel shows larger impact from the bulk ice in the downgoing direction. Overall this indicates that we can expect a final systematic uncertainty on the order of 10%

8.7.3 Combined Extremum Test

In this test we use the modified pseudo measurements as input and set all systematics to the value within the $\pm 1\sigma$ range where they have the maximum impact. We then perform the re-smearing test under these circumstances in order to investigate the balance point between bias and uncertainties. The results of this test showed a significant pull away from the central value at 10 iterations, which was found to be caused by the correlations between the systematic parameters not being correctly accounted for. These correlations were corrected for and implemented in future systematics tests.

8.8 Re-smearing Test with Modified Input

Here we seek to test the impact of our a-priori ignorance of the input spectrum. In order to assess the final number of iterations we must consider that we do not know the truth related to our measured

Hole Ice Systematics Impact

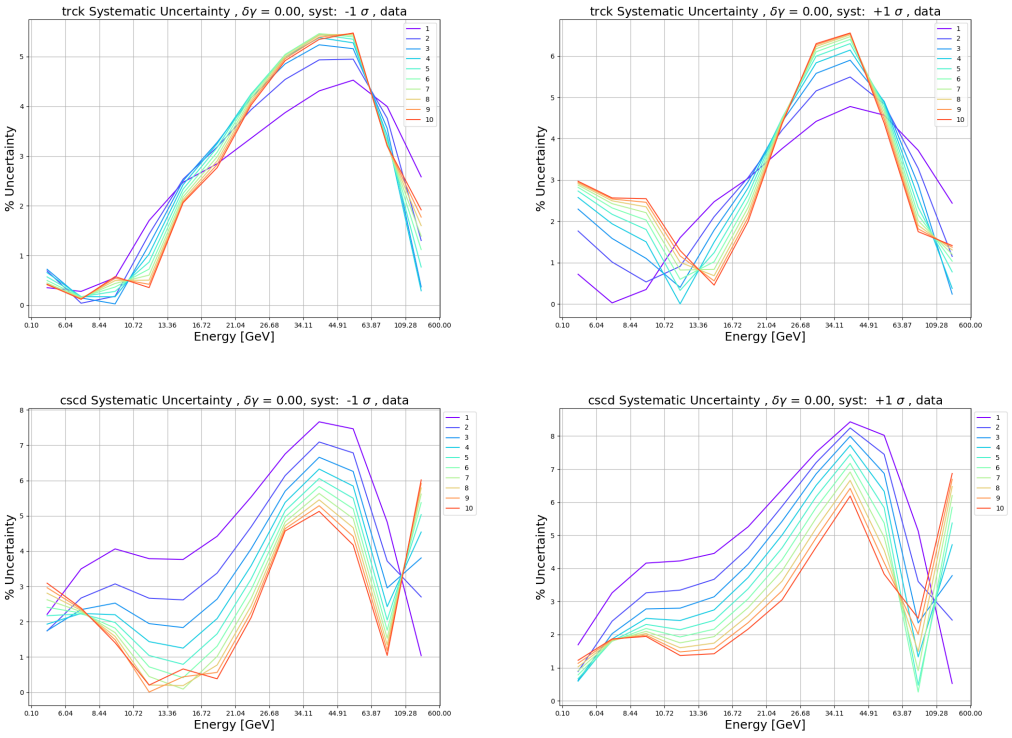


Figure 8.17: The relative impact of the Hole Ice in the re-smearing unfolding test. Top: track channel. Bottom: Cascade channel. Left: Hole Ice at -1σ. Right: Hole Ice at $+1\sigma$ i

Bulk Ice Scattering Systematics Impact

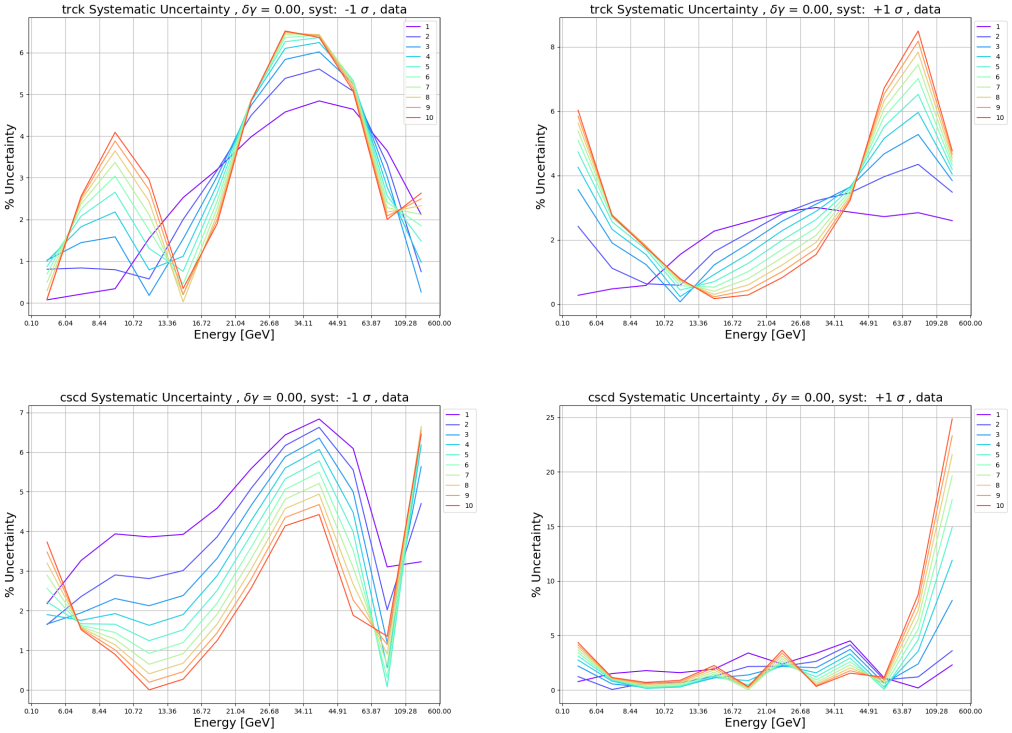


Figure 8.18: The relative impact of the Bulk Ice Scattering in the re-smearing unfolding test. Top: track channel. Bottom: Cascade channel. Left: Bulk Ice Scattering at -1σ. Right: Bulk Ice Scattering at $+1\sigma$

Bulk Ice Absorption Systematics Impact

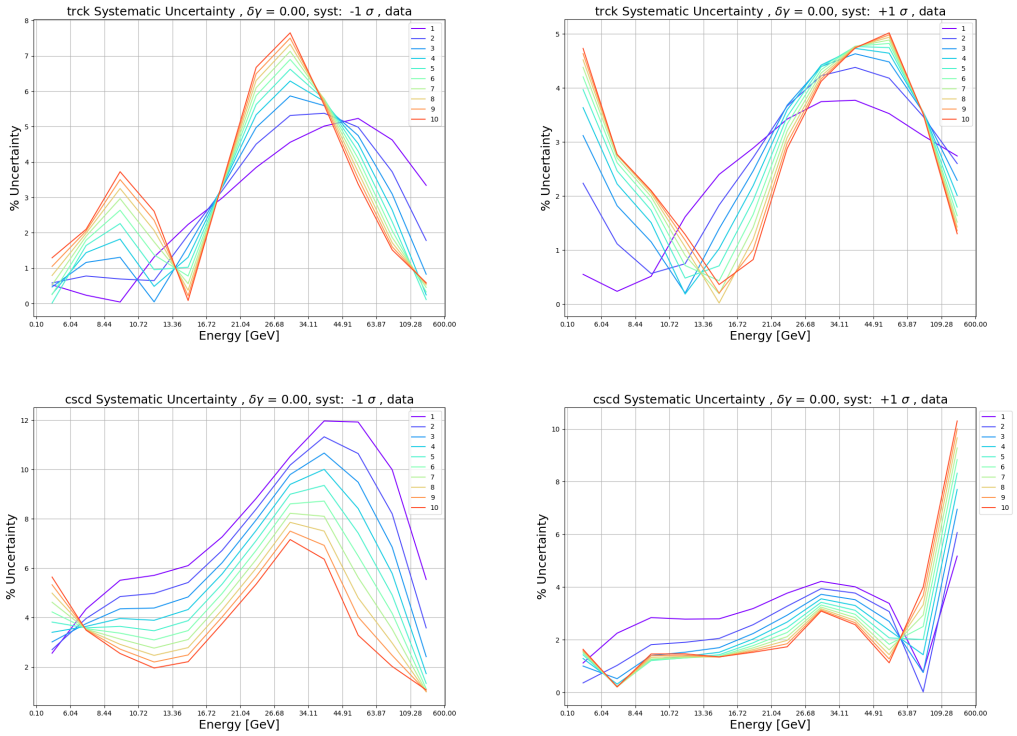


Figure 8.19: The relative impact of the Bulk Ice Absorption in the re-smearing unfolding test. Top: track channel. Bottom: Cascade channel. Left: Bulk Ice Absorption at -1σ. Right: Bulk Ice Absorption at $+1\sigma$

Atm. Index change	-0.15	-0.1	-0.05	0.0	0.05	0.1	0.15
Unfolded Events	74175.3	85818.9	100305.7	116923.9	136426.9	160018.2	188615.4
ν_μ normalization	0.85	0.9	0.95	1.0	1.05	1.10	1.15
Unfolded Events	102911.1	107214.6	112365.7	116923.9	120430.4	126201.5	130413.0

Table 8.2: Parameter variations in the re-smearing test for setting number of iterations.

spectrum - indeed this is what we are attempting to estimate. When making an actual measurement we do not know how systematically and statistically separated that measurement will be from the Monte Carlo used to generate the unfolding matrix. We see however, from previous tests, that when a systematic and statistical difference exists between the measurement and MC used, more iterations are needed to recover the true values (see sec. 8.3). All our previous tests have been performed with a known truth, the question now becomes how to optimize the number of iterations in case of an actual measurement with no known truth. We propose the following: we generate a pseudo-measurement by appropriately modifying the reconstructed MC - be it as a modification of the spectral index or normalizations of the different neutrino species. This pseudo-measurement will have no truth directly associated with it, and will in this test serve as our actual measurement. We now seek to investigate how well we can re-establish this spectrum through our unfolding and response matrix. As the measurement is statistically different from our MC the first step is to draw a number of pseudo experiments around the measurement, in order to generate an appropriate ensemble. Each pseudo experiment is then unfolded using the standard unfolding object, yielding an unfolded spectrum. In previous tests, this was then compared to the true MC spectrum, but as no such spectrum is available in this case, we instead do the following: we apply the response matrix to each unfolded spectrum and directly compare this re-smeared spectrum to the measurement from which it was generated. This will then, over an ensemble of 200 pseudo experiments inform us how many unfolding steps on average are needed in order to recover our measurement from an estimated truth. This method may seem cumbersome, but is necessary, since the difference from our actual measurement to our MC is inherently unknown in case of a blind analysis. We perform these re-smearing tests for a set of variations on the atmospheric spectral index as seen in table 8.2. The comparison between re-smeared and experiment is quantified the usual way via a test statistic, and we plot it averaged over the 200 pseudo experiments in the ensemble as a function of iterations. The test results are shown for the sampled values of the spectral index in fig. 8.20. The distribution indicates most values of spectral index shift have little to no impact on the analysis, however, larger shifts away from the nominal value clearly display the expected behaviour of a quick drop in test statistic. This shows the unfolding is converging rapidly on the input value, and indicates the unfolding procedure is able to compensate for the systematic difference in spectral shape between the input spectrum and the MC.

8.9 Closing Remarks on the Re-Smearing Test

With all the above considerations of the re-smearing closure test, a few comments are in order. The re-smearing test seems to be highly sensitive to systematic uncertainties and can only compensate moderately by the number of iterations, as high iteration counts lead to divergence. Furthermore, there is a logical consideration to be made: the re-smearing procedure in principle only tests whether the process of unfolding and then re-folding is considered closed (the process output converges on the input value). The unfolded result could fall anywhere within the parameter space, potentially very far from the true value, and as long as the re-smearing is closed, this test would not be able to reveal

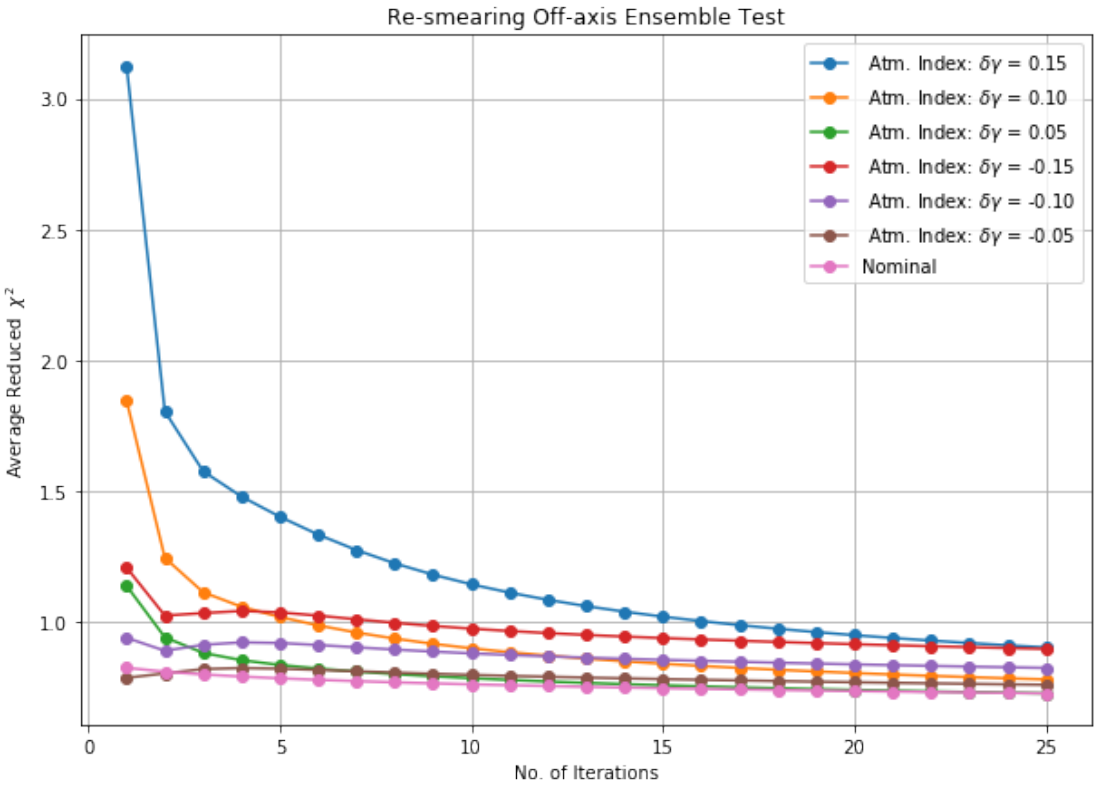


Figure 8.20: Average reduced χ^2 as a function of unfolding iterations. Each line represents a different ensemble of 200 trials drawn from our reconstructed MC modified by a factor on the spectral index. For all but the $\delta\gamma = +0.15$ case, the re-smeared spectra are within approximately 1σ of the pseudo experiment already after three to four iterations.

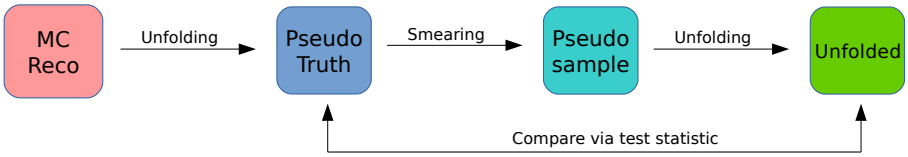


Figure 8.21: Analysis flow of the TSU closure test. A pseudo truth is generated by unfolding the reconstructed MC. This is then re-smeared, an ensemble is drawn and unfolded after which the final unfolded spectrum is compared the the pseudo truth.

it. Thus the re-smearing test does not reveal anything regarding the unfolded truth or say anything regarding how close that unfolded truth is to the actual truth of nature. Such a comparison is only possible when the truth is known, so is typically performed versus the best-estimate MC. Since the analysis method should be as model agnostic as possible, a method should instead be developed that is independent of the MC, but at the same time gives an indication of how close the unfolded spectrum resembles an underlying true spectrum. For this purpose, the truth-smearing-unfolding (TSU) test was devised, and it is the topic of the next section.

8.10 Closure Test: Truth-Smearing-Unfolding (TSU)

With the following test the aim is to show the unfolding is able to overcome a statistical and systematic difference between the natural truth (population) and the MC used to construct the response matrix. At the same time this test will account for the lack of prior knowledge of the natural truth, as this is a priori unknown in a real analysis. In previous sections the unfolding has been compared either to the simulation truth or re-smeared and compared to the input pseudo measurement. However, when running the analysis on real data it is impossible to compare to a truth in order to set a reasonable stopping condition. It is of course possible to compare to the MC truth, but that gives a strong bias towards the MC and basically voids the reasoning behind the entire unfolding procedure. It has also been shown that the variations due to re-smearing are small and that the re-smearing alone is not an optimal measure of the quality of the unfolding. Therefore the TSU-test was introduced to compensate for the a priori lack of knowledge of the underlying natural truth. The TSU test is shown in schematic form in fig. 8.21, where a pseudo truth is generated, smeared and then unfolded.

The pseudo truth is generated by unfolding the standard reconstructed MC using the standard 5 iterations as recommended by D'Agostini. The unfolded spectrum is then re-smeared by applying the response matrix (similarly to the re-smearing test). This re-smeared spectrum then forms a basis from which an ensemble of pseudo experiments are drawn and individually unfolded. In following the previous sections, the distributions involved are conceptually illustrated in fig. 8.22. The orange histogram represents the re-smeared standard unfolding, from which the pseudo data is drawn (black dots). Each is then unfolded (purple) and is tested via the usual test statistic against the pseudo-truth (green). This is done for 25 iteration steps and all results are recorded. The test statistic distribution is plotted as a function of number of iterations to set the stopping condition. 25 iterations was chosen as the maximum based on previous distributions showing minima at under 15 iterations, and to preserve computing power. This test reveals how many iterations we realistically need when the unfolded spectrum is statistically and systematically different from the truth used in creating the unfolding matrix. The next sections will cover the results of the TSU test in the statistics only case,

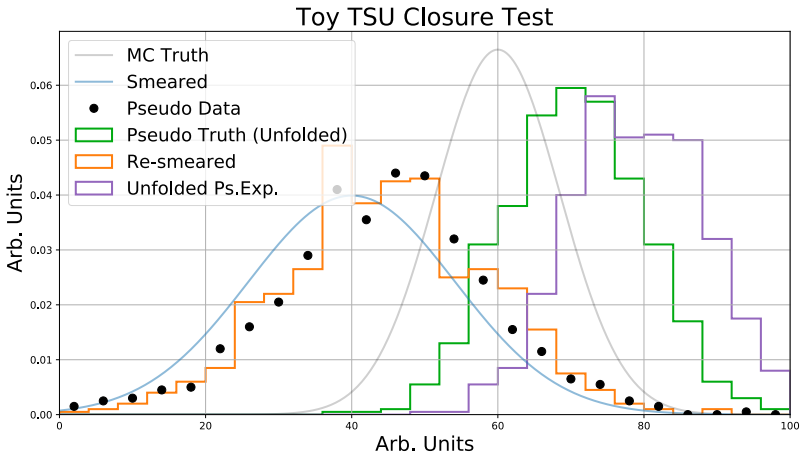


Figure 8.22: Conceptual Illustration of the distributions involved in the TSU test. A pseudo truth (green) is generated from the input MC by unfolding the reconstructed spectrum (blue) using the standard 5 iterations as prescribed by D'Agostini. This is then re-smeared (orange) and from the re-smeared distribution an ensemble of pseudo experiments are drawn (black dots). Each pseudo experiment is then unfolded (purple) and the stopping condition set by comparison between the unfolded and the pseudo truth.

as well as including systematics in order to prepare for a data challenge test.

8.11 TSU - Statistics Only

In order to confirm the impact on the analysis and assess the consistency of the TSU test, it is appropriate to run the TSU test in a statistics only case. All systematics are kept fixed at their nominal values and the pseudo measurement is taken from the standard MC. The unfolded results using 1 and 25 iteration steps are shown in fig. (8.23), and show good agreement between the unfolded spectrum and the MC truth, at the few percent level. In a realistic unfolding scenario we would look at the test statistic distribution to set the stopping condition, which brings us to the next section.

8.12 TSU Systematics Test

Having a good understanding of the statistical behaviour in the TSU test we can now proceed to include all the systematic parameters. This includes the discrete systematics described in chapter 7, both on the expected spectra and the selection efficiency. The implemented method uses the standard bracketing approach for systematics described in section 7.1. As the systematic uncertainties impact the calculation of the unfolding matrix, they are expected to have a significant impact on the unfolding itself. The Bayesian iterative unfolding can compensate to some extent: with a high number of iterations the unfolding can enter into the divergent regime. The results of the TSU test including systematics is shown in fig. (8.25), and we can conclude:

- Systematics impact the unfolding

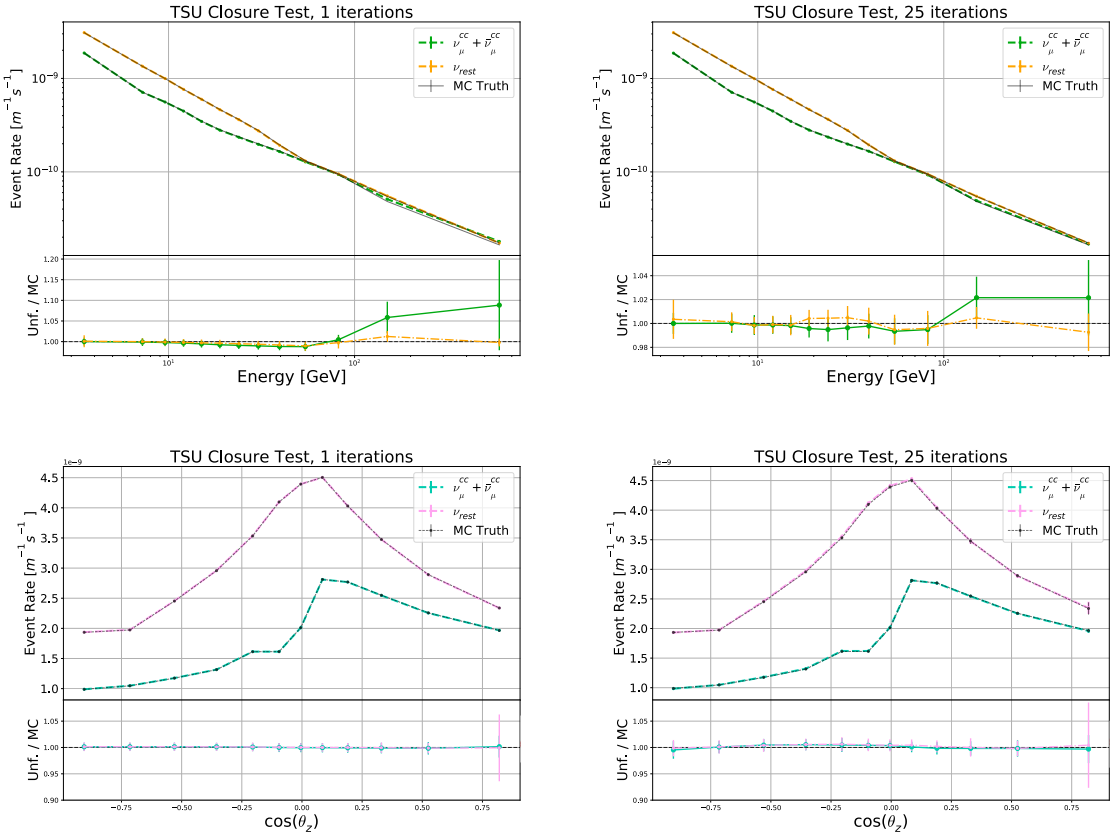


Figure 8.23: Unfolded pseudo experiment in the statistics only case, for 1 iteration (left) and 25 iterations (right). Top row shows energy spectra, bottom row shows the zenith angle dependency.

- Bayesian unfolding can compensate to some extent
- Impact from systematic depends on number of iterations

In continuation of these points, it is appropriate to turn our attention to the test statistic scores as a function of iterations. As shown in fig. (8.26), the choice of test statistic now plays an important role in setting the stopping condition. When using the ROOT χ^2 definition (eq. 8.7), which ignores uncertainties the unfolding consistency is seen to increase with the number of iterations. This however measures only the distance between the data points, while also considering the integrated spectrum. The consistency increasing over iterations is illustrated in fig. 8.24, showing the TSU MC Truth and the unfolded spectrum for several iterations. A clear convergence towards the truth spectrum with the number of iterations is observed. When also considering statistical and systematic uncertainties as in 8.9, the picture changes: The most favorable iteration is now one, as the unfolding converges well within the error bars after the first iteration. While the difference between unfolded spectrum and the pseudo truth decrease, the error bars on the unfolding on the other hand increase and dominate over the former, leading to an increase in test statistic. The valid choice is to include the total uncertainty, since the iteration-to-iteration variation in the unfolded spectrum is smaller than the uncertainty and this method includes more relevant information. Section (8.6) and fig. (8.10) prescribes how to set the stopping condition in case of the test statistic distribution seen in the right panel in fig. (8.26): Ignore the first four iterations to avoid bias, and then minimize the distribution from there. In this case such a method leads to setting the stopping condition to 16 iterations. With the considerations on systematics and how to set the stopping condition thoroughly tested on simulation, the next section describes similar testing performed on a small subset of the final data sample.

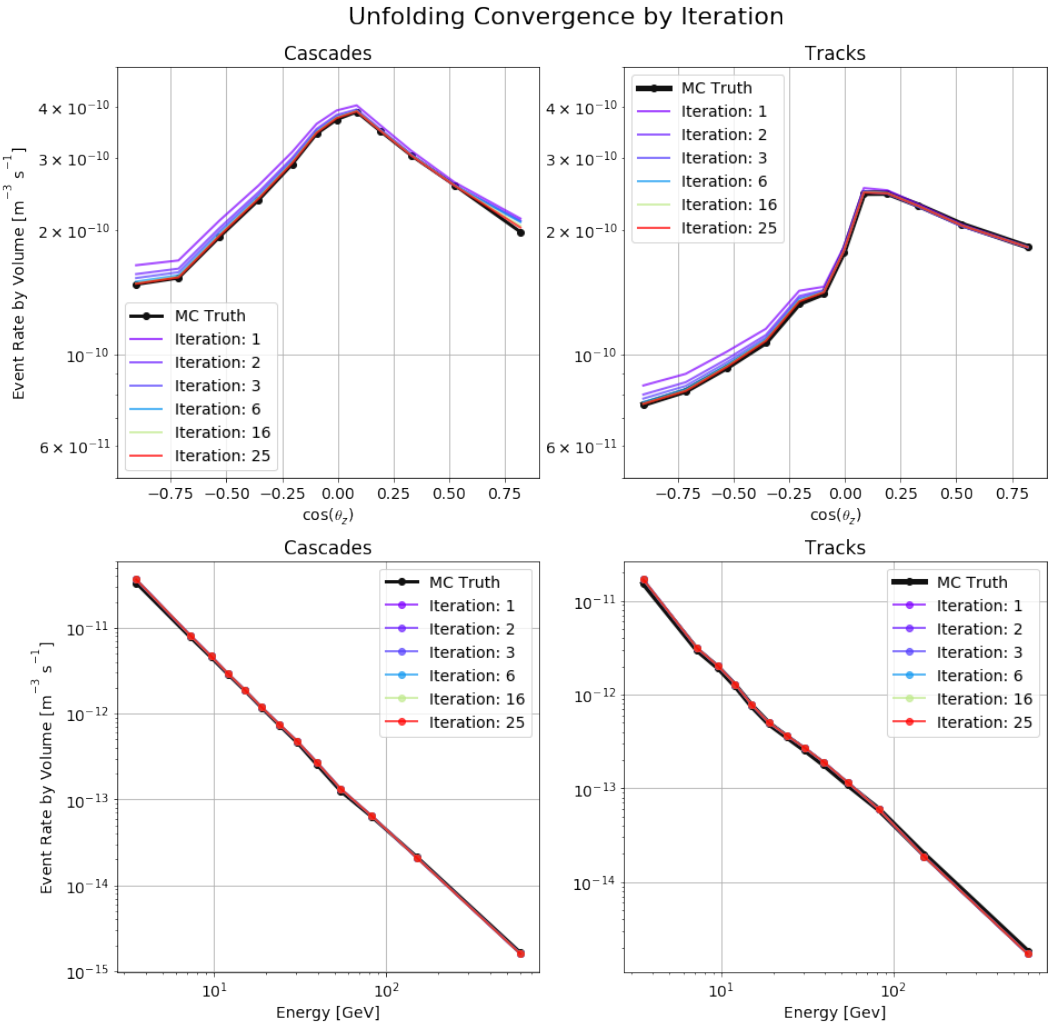


Figure 8.24: True and unfolded distributions for the TSU unfolding test. Top row: zenith angle spectra. Bottom row: Energy spectra. Left: Cascade channel. Right: Track channel. The unfolded spectrum converges on the TSU truth with the number of iterations in all insets, but can be difficult to identify for the energy plots due to the power law behaviour of the spectrum.

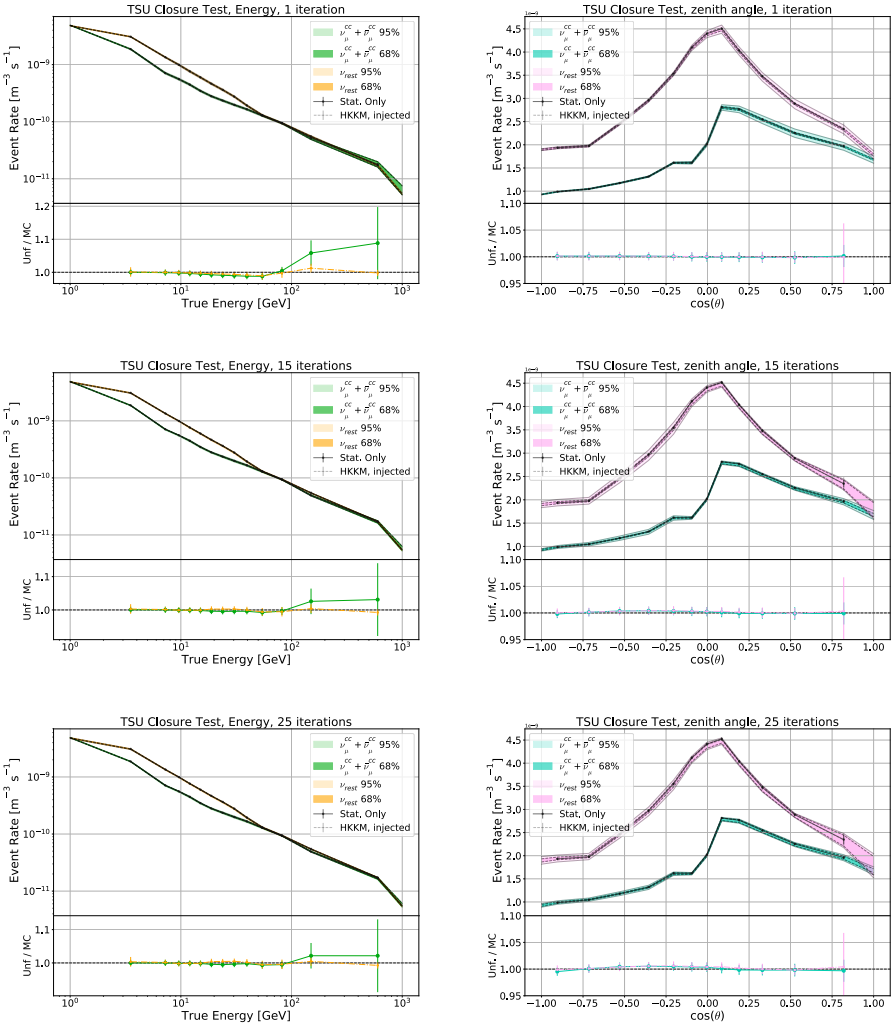


Figure 8.25: Marginalized unfolded distributions for the TSU closure test including systematics. Left column: Energy. Right column: $\cos(\theta_z)$. Top: 1 iteration. Middle: 15 iterations. Bottom: 25 iterations. The highest energy bin slowly converges on the input MC with the number of iterations, where as the upgoing and horizon regions are being pulled away from the MC with the number of iterations. The stopping condition is calculated using the unmarginalized 2D spectrum.

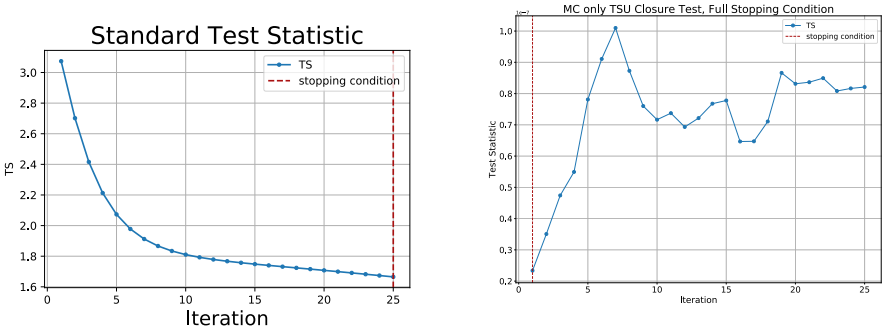


Figure 8.26: TSU closure test stopping conditions. Left: Standard calculation based on eq. 8.7. Right: Full test statistic based on minimization of the total uncertainty as given in eq. 8.9. The contrast between the two is striking, with the full test requiring only one iteration as the unfolding is consistent within the error bars already at that stage.

8.13 Burn Sample Unfolding

In order to test the analysis method on real data without unblinding the analysis, a small subset of data was prepared: 10% of the full sample was used, with a random sampling from each of the five analysis years. This was then used as input for the TSU test, in order to check consistency with the MC only tests, before proceeding to blind checks on the full data sample. Two burn sample tests were performed: One using the standard burn sample and one in which all events were weighted by a factor of ten to compensate for the lack of statistics. The test statistic stopping condition for the two cases are shown in fig (8.27). The two distributions are clearly different, illustrating the impact on the stopping condition a lack of statistics can have in this analysis. At first glance the standard burn sample seems to have lower test statistic values than the scaled version. However, it is important to realize the test statistic minimizes the total uncertainty band on an absolute scale, leaving the normalizations not directly comparable. Any effect on the stopping condition directly transfers to the unfolded results, underlining the need for adequate statistics. Moving on to the unfolded results from the burn sample tests, the statistics only case is shown in fig. (8.28) (unscaled) for 2 and 21 iterations, in order to illustrate the state of the unfolding in the unscaled case. The full uncertainty treatment is shown in fig. (8.29) (unscaled) and (8.30) (scaled). Both the unscaled and the scaled versions are included to illustrate the impact of sample size on the overall uncertainty. at 21 iterations the statistics only case shows good agreement to within a few percent with the MC Truth, which in this case is represented by the input pseudo truth and can thus be different from the standard MC truth.

The results of the burn sample test are mostly consistent, however, a few discrepancies were discovered, which will be discussed in the following sections.

8.13.1 TS Distribution - checks on width of distributions

The 200 individual trials drawn from the re-smeared spectrum are unfolded and compared against the pseudo-truth using the χ^2 test given in eq. 8.7. The average value of the χ^2s is plotted as a function of iterations and is shown in fig. 8.31. The error bars show the square root of the variance of the test statistic distribution for the ensemble. This however, assumes an underlying Gaussian

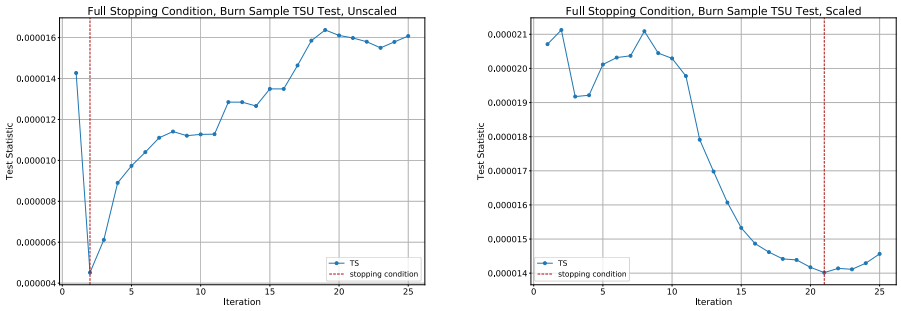


Figure 8.27: Stopping condition, showing the average TS of the ensemble as function of iterations, with the minimum marked in the dashed red line. TS is calculated as given in eq. 8.11. Left: Standard 10% burn sample. Right: Scaled burn sample. At first glance the standard burn sample seems to have lower TS values than the scaled version. However, it is important to realize the test statistic minimizes the total uncertainty band on an absolute scale, leaving the normalizations not directly comparable.

distribution. Interestingly the error bars in the MC test grow with the number of iterations up to about ~10 iterations, before declining as iterations increase. This behaviour has not been seen in any of the previous tests and is unexpected. The size of the error bars as function of iterations is shown in the right panel of fig. 8.31. A histogram of pseudo experiments in the ensemble as a function of their test statistic and grouped by iterations is shown in fig. (8.32). The widening of the error bars is then understood to be a consequence of the individual pseudo experiments in the ensemble converging towards the truth at slightly different rates. Not only does the central value of the TS distribution shift to lower values with the number of iterations, the shape of the distribution also changes, yielding the unexpected behaviour in the right panel of fig. (8.31).

8.14 Burn Sample: An Unfolding Discrepancy

This section will take a look at a historical version of the TSU test utilizing a 10% burn sample, which led to an update of the binning used in the analysis. These were preliminary tests and revealed a few issues which had to be corrected. Investigating the marginalized spectra for the unfolded burn sample in fig. 8.33, revealed two issues: firstly the highest energy bins display a bias upwards of 10% as well as a large uncertainty compared to the MC-only TSU test (which has uncertainty on the one percent level, see fig. 8.25). The discrepancy for the marginalized energy spectrum is illustrated in fig. 8.34, which shows the ratio of the unfolded spectrum to the MC truth. Secondly the ν_μ^{cc} channel shows a large systematics impact of up to 20% in the mid-range energy bins as can be seen in figures 8.33 and 8.30. These effects could be caused by events creeping into the analysis region from outside the binning range. A series of checks were implemented in order to further investigate the discrepancies. In the following sections these checks and the resulting outcomes are discussed.

8.14.1 Binning Optimization Revisited

A possible cause of the discrepancy in the unfolding outcome was suspected to be data/MC mismatch, which has been known to be an issue in other analyses utilizing the GRECO sample. In those cases

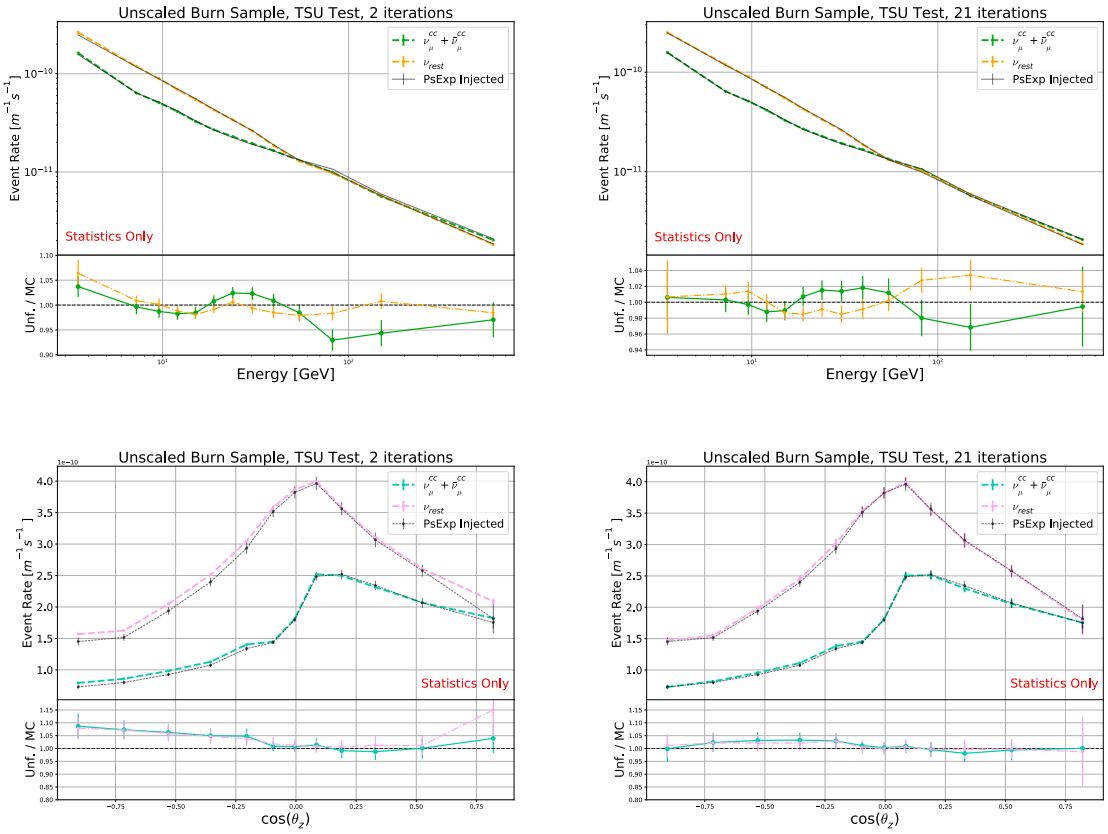


Figure 8.28: Burn sample unfolding in the statistics only case.

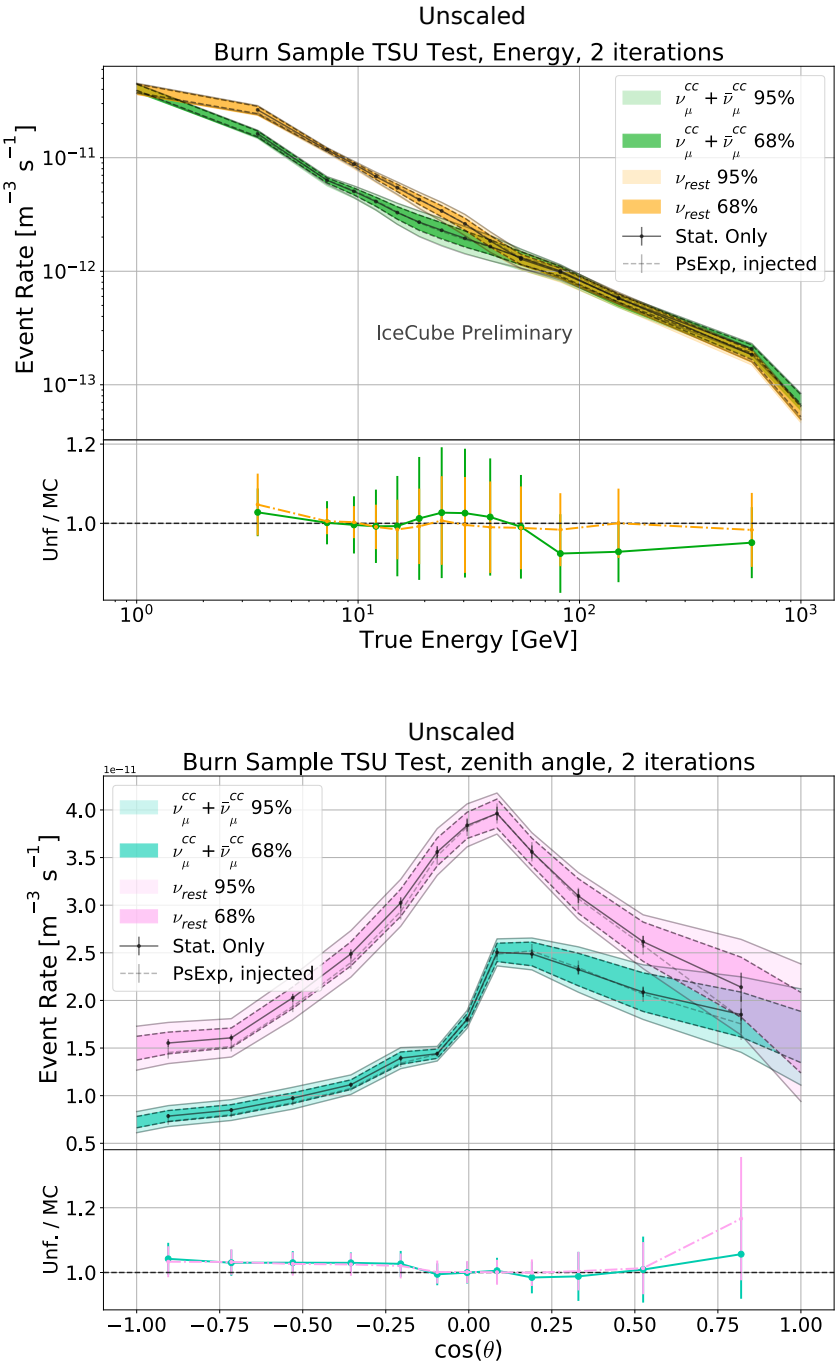


Figure 8.29: Burn sample unfolding in the full systematics case.

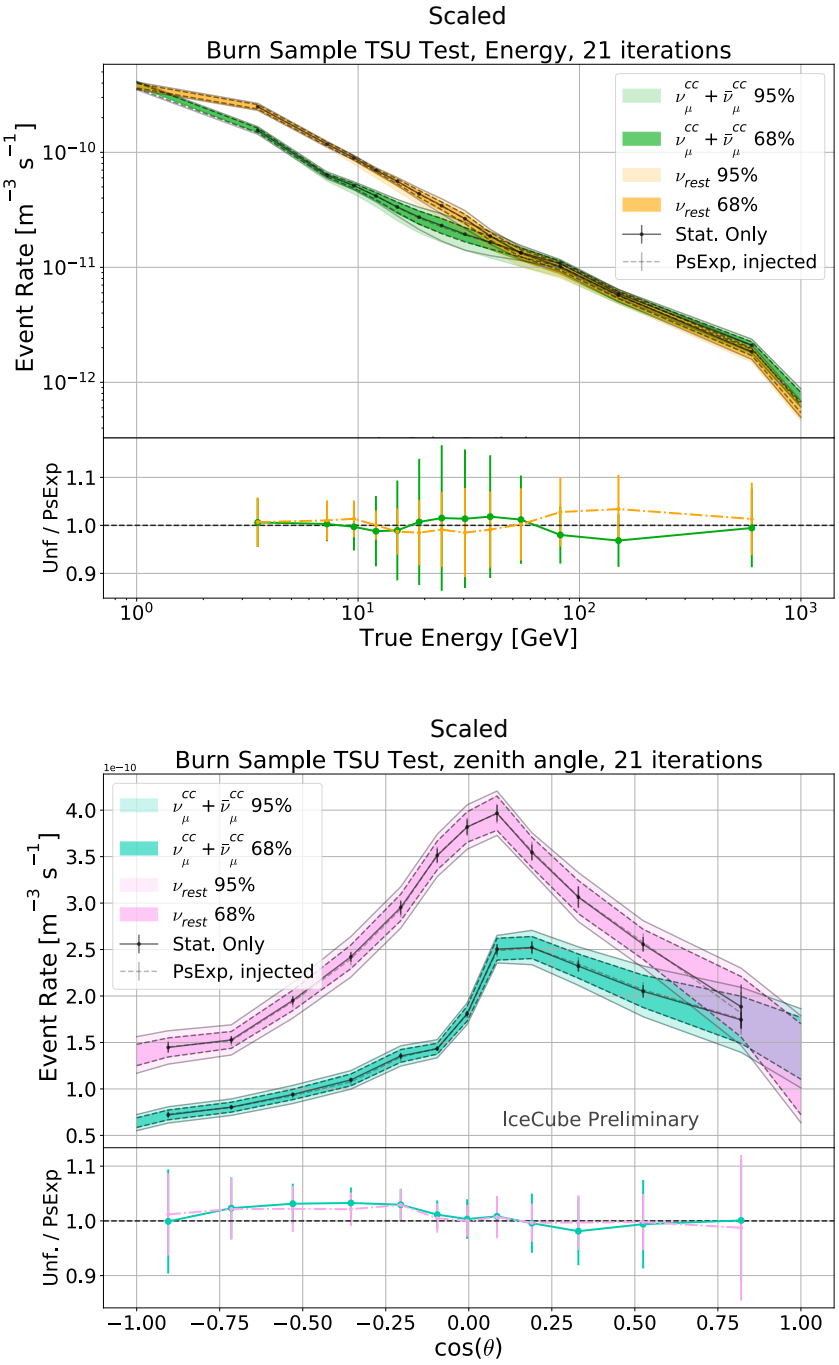


Figure 8.30: Burn sample unfolding in the full systematics case.

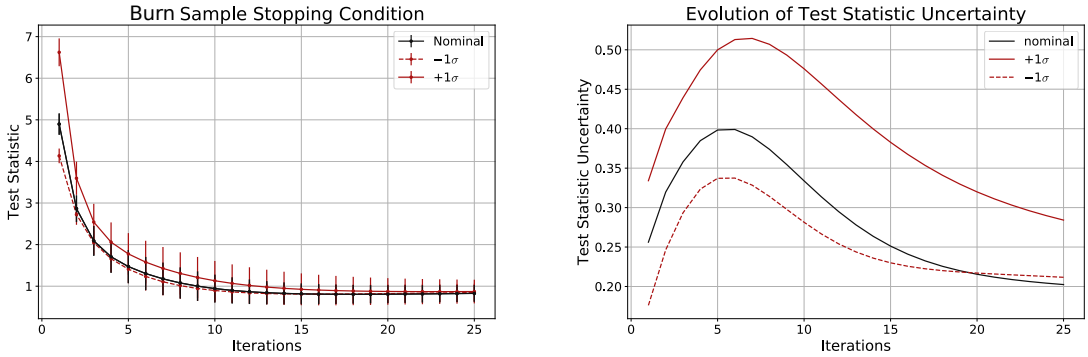


Figure 8.31: Left: Average χ^2 as a function of iteration. The error bars represent the square root of the variance in the ensemble. Right: Error bar size as function of iterations. Interestingly the error bars grow up to ~6 iterations, before dropping off. The slight cross-over at high iteration number is due to the behaviour of the unfolding matrix, when having less statistics.

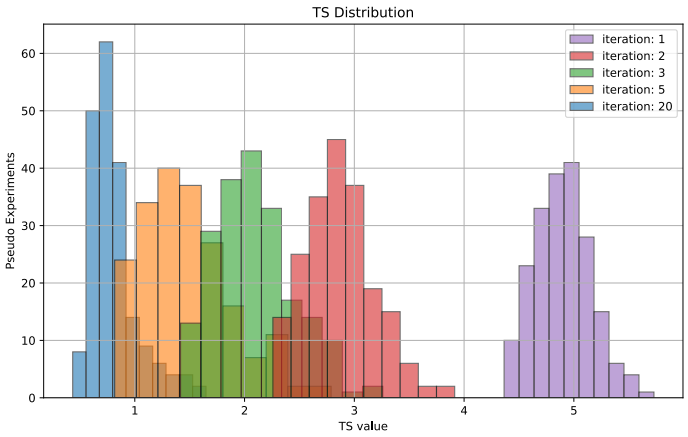


Figure 8.32: Distribution of TS values for the ensemble, grouped by iteration. As iterations increase the TS decreases and the distribution slims considerable (albeit with a bit of an upper tail). Distributions for intermediate iterations values are seen to be wider.

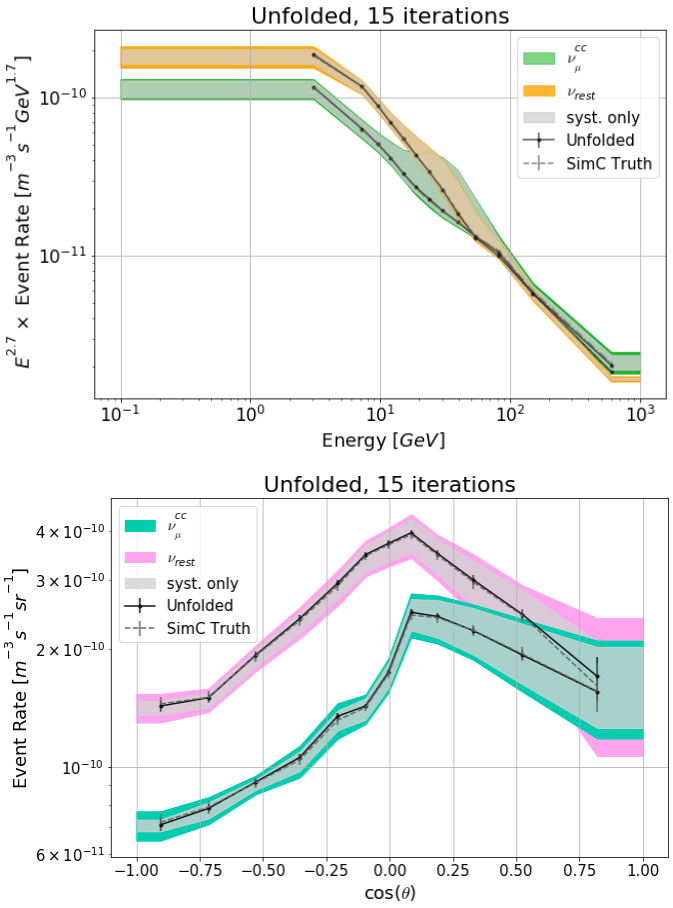


Figure 8.33: Unfolded results of the burn sample TSU test - Left: Energy. Right: $\cos(\theta_z)$. Notice the width of the uncertainty band for the energy plot widening in the energy range between 10 and 100 GeV.

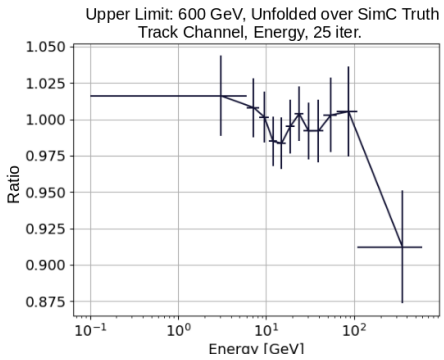


Figure 8.34: Outcome of the burn sample TSU closure test for $\overset{(-)}{\nu}_{\mu}^{CC}$ channel using the standard binning, reaching up to 300 GeV in reconstructed space and 600 GeV in true space. Graph shows the average outcome of the ensemble relative to the pseudo-truth used in the TSU test. As can be seen the highest energy bin suffers from poor consistency and high variance.

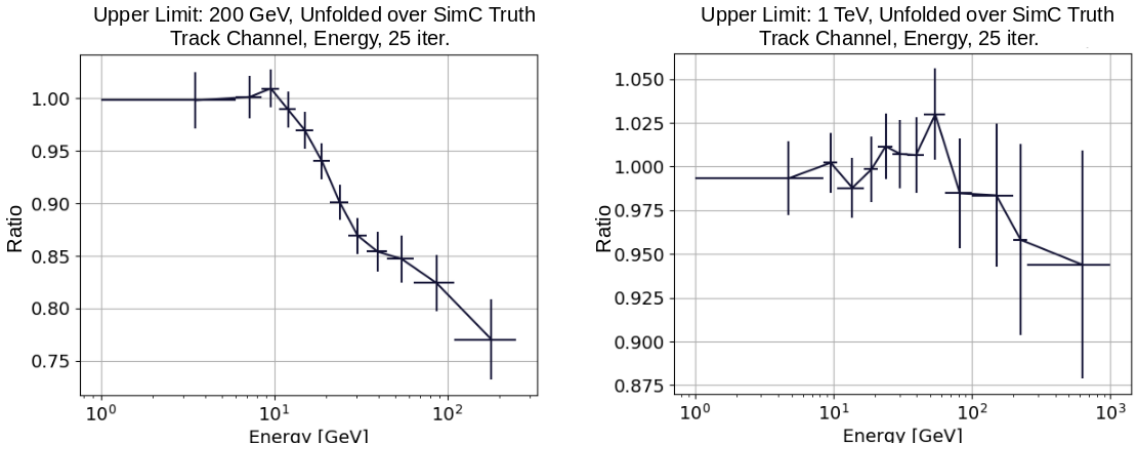


Figure 8.35: Outcome of the TSU test when changing the binning. Left: Lowering the upper edge of the highest energy bin from 600 GeV to 250 GeV, a large discrepancy over almost the entire energy range can be observed. Right: Adding two more bins with edges at 200 GeV, 250 GeV and 1Tev shows the unfolding to be stable up to about 200 GeV, but does not resolve the discrepancy at higher energies.

the discrepancies were solved by lowering the energy threshold and imposing a cut at $E < 54$GeV [60]. Building on those previous observations, the first attempt was to lower the edge of the highest energy bin from 600 GeV to 250 GeV. The TSU test was then repeated and the result is shown in the left panel of fig. 8.35. This led to an increase in the discrepancy and a further shape difference across almost the entire energy spectrum as compared to the results using the standard binning (fig. 8.34). This is a clear indication the effect is different from what was seen in other GRECO analyses. Extending the binning range to 1 TeV and adding another two bin edges at 200 GeV and 250 GeV revealed the unfolding to be relatively stable up to about 200 GeV as can be seen in the right panel of fig. 8.35. However, extending the binning up to 1 TeV did not fully alleviate the issue. This also gave an indication that the discrepancy was only present at high energies, and a possible cause could be migration of events from outside the binning range.

Plotting the true energy distribution of events for each reconstructed energy bin illustrates the bin coverage of the sample and is shown in fig. 8.36. As can be seen from the final cell at the bottom right, even with an energy cut at 1 TeV, the sample is still under represented in the highest energy reconstructed bin. This underrepresentation due to the highest bin edge is highlighted in fig. 8.37. Many events are lost and neither included in the response matrix nor in the unfolding matrix. Thus these events create an under fluctuation in the estimates of the detector performance, leading to the discrepancy seen at energies above 200GeV. Adding yet another bin from one to five TeV to catch this in the unfolded spectrum solves the issue as shown in fig. 8.38. However, the highest energy bin shows large uncertainty and there is no way to further verify that all the possible migration effects have been covered. This bin is therefore treated as an 'overflow' bin which only exists to catch these events and help provide the correct coverage and stability for the rest of the unfolding bins. The content of this bin is not included in the final analysis, and the updated binning is shown in table 6.2. The updated burn sample TSU test with the revised binning is shown in fig. 8.39.

8.14.2 Muon Background Investigation

A high level of impact from systematic uncertainties are seen in the central energy bins between 10GeV and 100 GeV, and while somewhat alleviated still persist after the binning correction in the previous section. At first glance this could be an artifact of the unfolding treating the muon background incorrectly, with events migrating into the sample from outside the energy range. However, with the changes above, the unfolding now goes all the way up to five TeV (including the overflow bin), adequately covering the entirety of the energy region of the muon component.

The high systematic uncertainty seen in the central energy bins was suspected to be caused by background. The true and reconstructed MC muon distributions for energy and zenith angle are shown in fig. 8.40. The main Muon contribution in reconstructed space coincides almost one to one to the location of the unexpected uncertainty in the energy spectrum. The impact on the overall uncertainty is thus likely to arise from the number of muons reconstructed into those bins, which are then unfolded into the background bin. The standard Muon suppression in the GRECO sample, even though highly efficient, is not able to remove enough muons that this effect is not seen, hinting at a discrepancy between the MC and data for the muons. A thorough investigation regarding data/MC agreement of the muon background could be useful, but is beyond the scope of this thesis at this point. However, as illustrated in fig. 8.40, about 70% of all reconstructed muons fall in the downgoing region above the horizon. In order to suppress the muon background in the marginalized energy spectrum, a cut on the zenith angle is enforced. This will of course also affect neutrino event rates in the downgoing region, however, since most neutrino events are either upgoing or around the horizon it is reasonable to

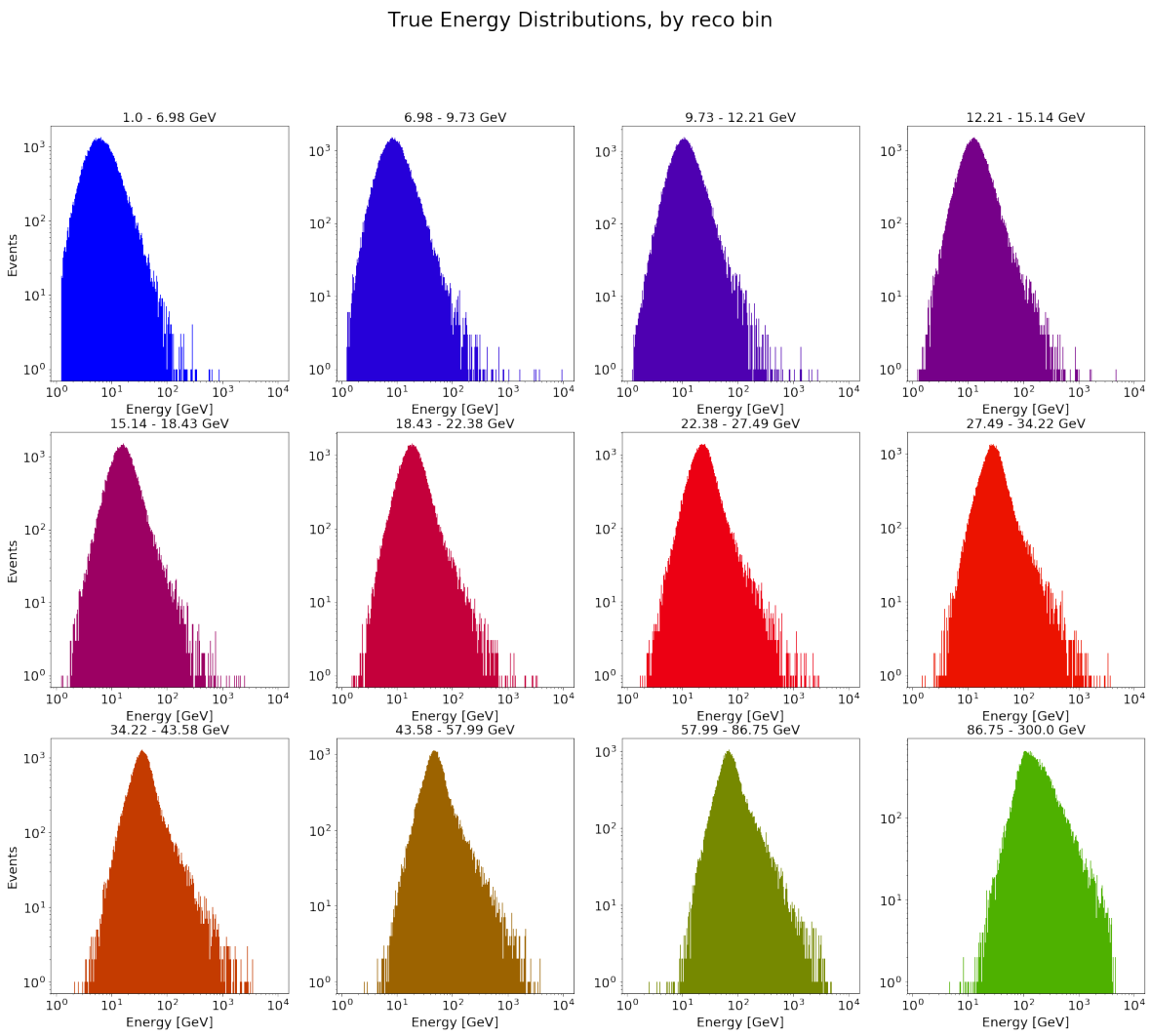


Figure 8.36: True neutrino energy by reconstructed bin, giving an overview of which true energy values end up migrating into which reconstructed bins. Events are shown unweighted, in order to illustrate the spectrum used for building the response matrix. The standard binning goes up to 600 GeV. The sharp line in the top left panel is due to the IceCube simulation not extending below 1 GeV.

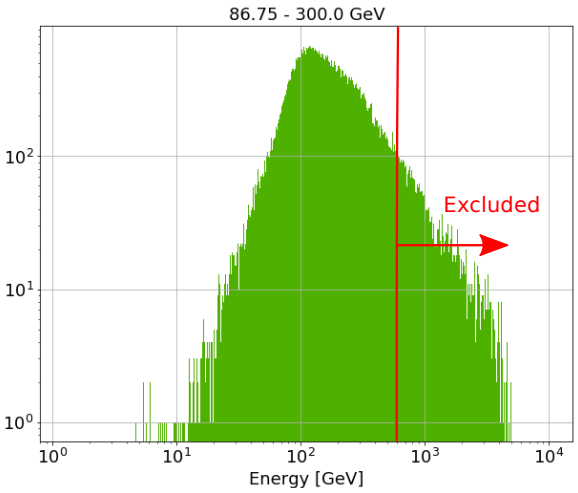


Figure 8.37: True energy coverage for the final reconstructed bin from 100 GeV to 300 GeV. Because the standard true energy binning only goes up to 600GeV in true space, many events are lost and not included in the unfolding matrix.

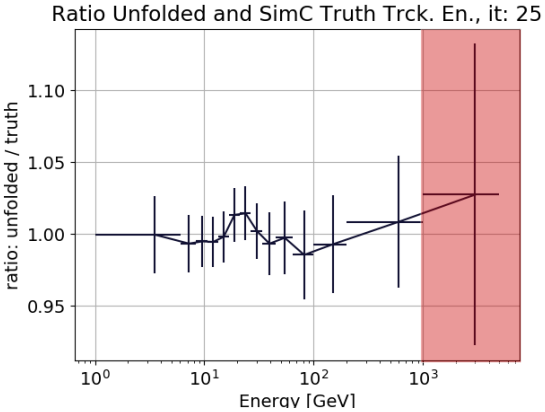


Figure 8.38: TSU unfolding closure test after application of fixes in section 8.14.1 and 8.14.2. Data points show the average unfolded value of the 200 pseudo experiments. Two bins with edges at 200 GeV, 1 TeV, and an upper edge of 5 TeV were added. The final bin marked in red is treated as an overflow bin and not included in the final analysis.. The unfolding is consistent to within a few percent up to 1TeV. The highest energy bin from 1-5 TeV, marked in red, is treated as an overflow bin and not included in the analysis.

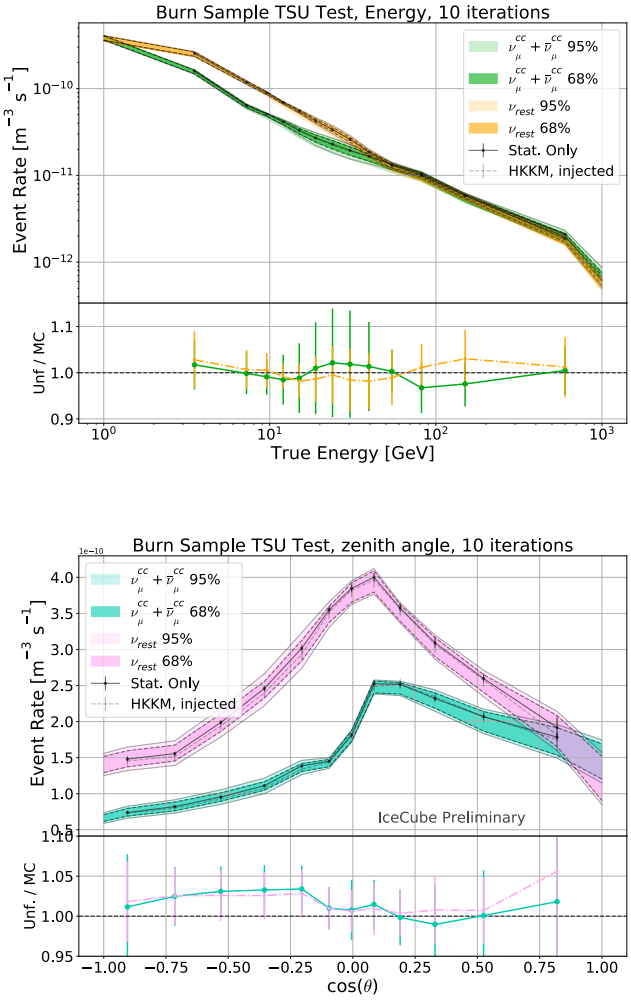


Figure 8.39: Burn sample TSU test performed with updated binning. The last bin bias has been reduced to about 5% and now falls within one standard deviation of the pseudo truth.

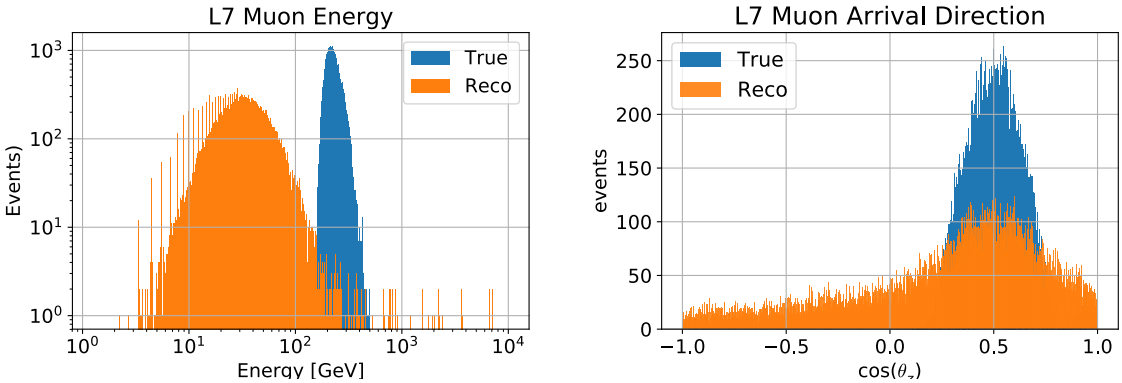


Figure 8.40: Distribution of muon events at final level. Left: Energy. Right: Cos(θ_z). As expected no upgoing muons are seen.

apply a cut at $cos(\theta_z) > 0.04$. Note this cut is only enforced when marginalizing the unfolded spectra, and thus has no effect on the unfolding matrix.

8.14.3 Resolving the Discrepancy

The fixes described in the two previous sub-sections are implemented in the TSU-test. The consistency plot in fig. 8.38 reveals the issue to be largely resolved. The unfolded results using the updated binning and the new zenith angle cut are illustrated in fig. 8.42. The analysis is rerun and the stopping condition plot is shown in fig. 8.41, giving the optimum number of iterations as 21, which is consistent with the previous scaled burn sample test. The results for the statistics only case, as well as the full systematics treatment using the sampling approach from the following section are shown in fig. 8.42 and fig. 8.43 respectively. The updated burn sample TSU test displays a far more reasonable impact from systematics in the region of the energy spectrum previously afflicted with large systematic errors, and the bias in the highest energy bin has been resolved. This means the bias is so small it can be safely ignored, since all data points agree with the truth well within the standard 1σ error bars.

8.14.4 Systematics: Bracketing vs. Random Sampling

The preliminary method selected for this analysis was a bracketing approach due to historical reasons. However, further investigation of the systematic uncertainties, particularly the oscillation parameters, led to the conclusion that not all parameters had their maximum pull values at their 1σ extent of their prior. Another approach was necessary to investigate the over/under estimation due to this discrepancy. Drawing trial sets out of all systematic parameters within their priors has previously been used in literature [3], but runs the risk of under-sampling the systematic parameter space if the number of samples are smaller than ~10 samples per dimension. In our case we have shown certain parameters to be dominant while others have only minimal effects on the expected spectrum shape, as discussed in chapter 7 and section 8.7.2. The systematic impact in this analysis is dominated by the four discrete systematic parameters, meaning it is feasible to draw random samples from our parameter set and unfold those in order to get an estimate of the uncertainty. 1474 trials are drawn and each trial is taken as pseudo-data as input for the TSU-test. As random trials are drawn each systematic is

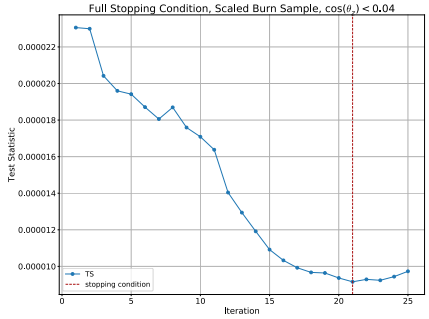


Figure 8.41: Test statistic as function of iterations for the scaled burn sample unfolding including the updated binning and the $\cos(\theta_z) < 0.04$ cut. The optimum number of iterations is 21, which is consistent with the scaled burn sample test without the modifications.

sampled at 1474 random values and as only 4 systematics dominate, this should be an adequate number of samples to get a good estimate of the uncertainty. The unfolded spectra are shown in fig. 8.44 as gray curves superimposed over the previous result from the bracketing approach. Generally the two approaches show good agreement, however, the bracketing approach seems to be overly conservative in certain regions where it over estimates the uncertainty compared to the random sample approach. This is expected, as the bracketing approach is known to utilize a box in the systematic multidimensional space, where a more realistic approach would give a multi dimensional ellipse. With this random sampling approach the uncertainty bands can be calculated as the 68% and 95% quantiles over the unfolded trial ensemble. This was done both for the standard MC TSU-test and the TSU-test on the burn sample. The updated plots are shown in fig. 8.43. This method has been selected as the approach for handling systematics in the final analysis unfolding.

8.15 Data Sample TSU Test

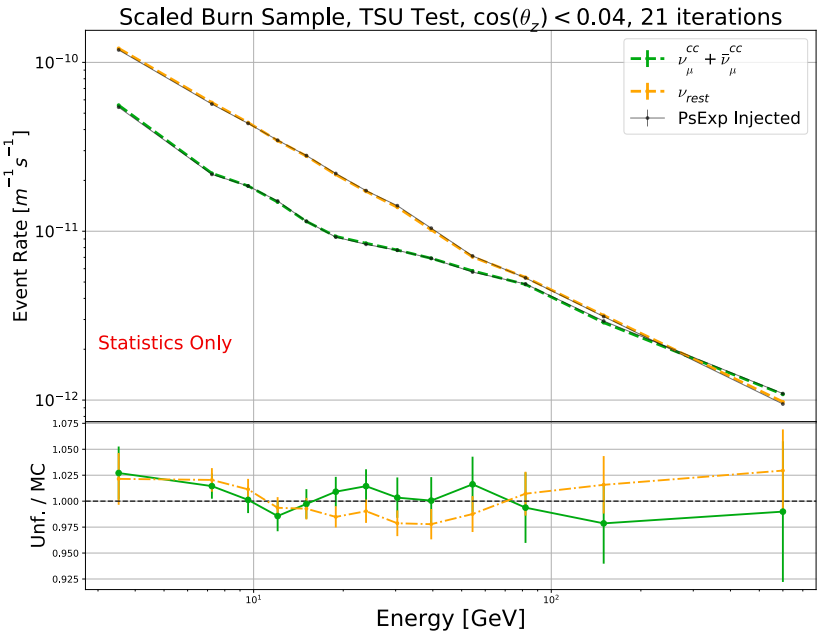
Using the TSU test (sec. 8.10), with the random sampling method (sec. 8.14.4), to estimate the systematic uncertainties, the full data sample is unfolded using ~1600 samples. The pseudo experiment truth is again set by unfolding with 4 iterations and drawing a Poisson distributed pseudo experiment around the unfolding. The following blind checks were in place before unblinding:

1. Test statistic, stat. only: convergence is expected, in consistency with the outputs from the burn sample tests.

2. Test Statistic, full: the test statistic is calculated from eq. (8.9). The distribution is expected to take on a similar shape to the distribution from the MC-only TSU test, with a quick drop in TS value and a clear minimum. Differences could indicate growth of either statistical fluctuations, enhancement of unexpected spectral features or a large systematic impact.

3. Unfolded Ratio: the unfolded spectrum should lie within a few percent of the injected pseudo experiment truth. This however, needs to be tested as a ratio because blindness is to be conserved.

A post unblinding check was also introduced in order to check for clear bias in the unfolding:



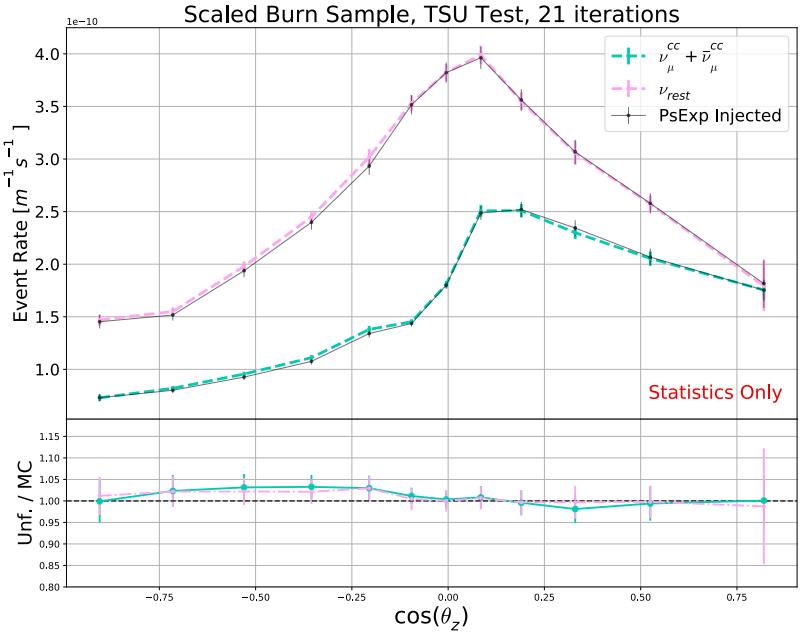


Figure 8.42: Statistics only result of the scaled burn sample unfolding using the updated binning and cos(θ_z) cut, discussed in sections 8.14.1 and 8.14.2 respectively. With the two fixes implemented the unfolding falls within 2.5% of the injected truth even at 25 iterations.

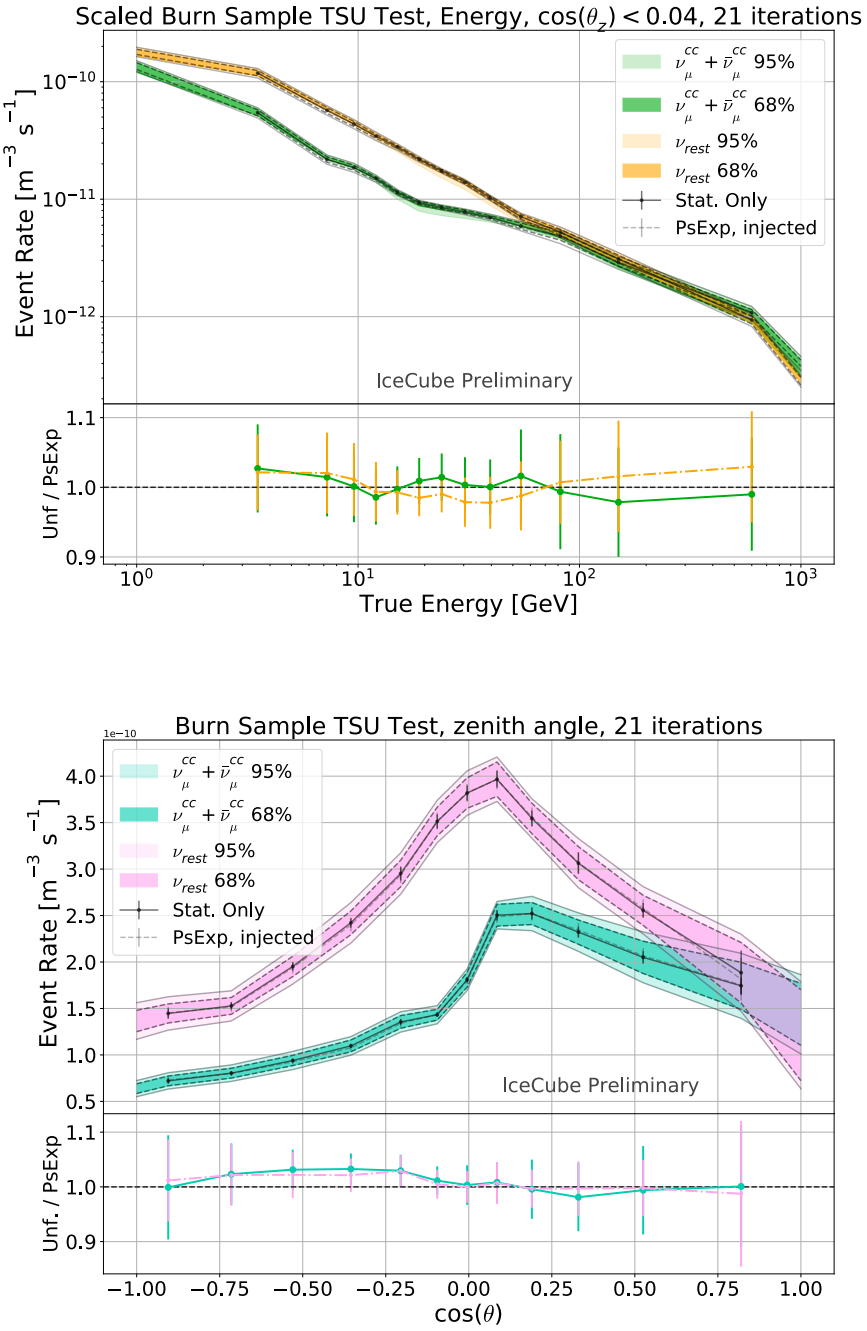


Figure 8.43: Scaled burn sample unfolding using the updated binning and cos(θ_z) cut, discussed in sections 8.14.1 and 8.14.2 respectively. Here the full uncertainty treatment discussed in section 8.14.4 is included.

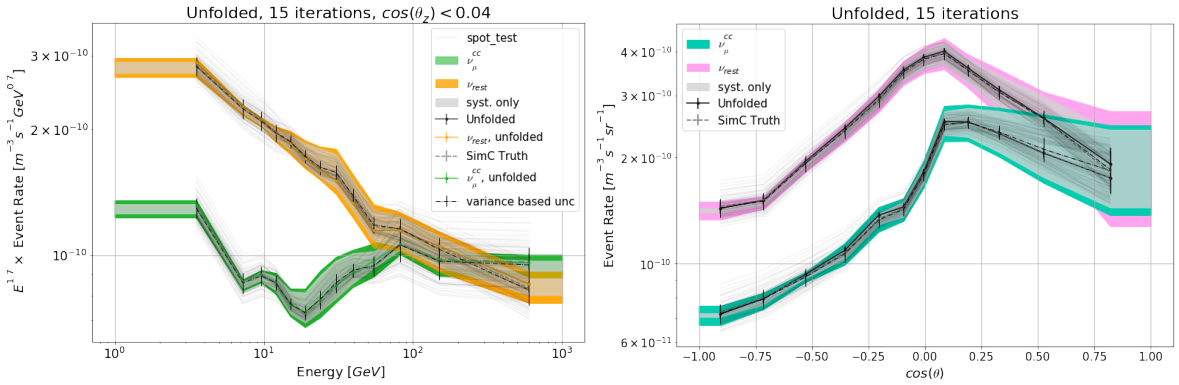


Figure 8.44: Brazilian 'hair' plots - standard bracketed unfolding with the unfolded trials of the random sampling superimposed in gray.

- Unfolded shape: is a-priori expected to be different from the GRECO MC. This can be checked via the ratio between the unfolded values and the MC, while keeping the scale on the y-axis unknown, in order to preserve blindness. In case the unfolded histogram matches the MC well within one standard deviation this could be indication of a bias in the unfolding method. It could also indicate the MC to be a precise representation of nature, however, this is a-priori unexpected. There is no obvious method for distinguishing the two cases, but a clear 1:1 consistency should be investigated more carefully.

The unfolding was performed blind via the TSU test. The stopping condition was calculated and is illustrated in fig. 8.45, setting the stopping condition at 17 iterations. For comparison the burn sample stopping condition came to 21 iterations. The unfolded ratio to the injected pseudo experiment truth is shown in fig. 8.46. All unfolded data falls within 4% percent of the injected pseudo experiment, and agree within one standard deviation. This gives a good indication the third unblinding condition is fulfilled. The analysis is unblinded and the unfolded results for the TSU test are shown in fig. 8.47. Notice this is not the same as final data unblinding, as the TSU test unfolds the re-smeared injected pseudo experiment. The results fall on the expected scale, and show relative uncertainties within a few percent. These are illustrated in fig. 8.48 and a short discussion is in order: for the energy spectrum the largest relative uncertainties come to 10% with the central energy bins showing markedly less uncertainty down to about 5%. For the zenith angle distributions the results are far more precise, yielding a relative uncertainty in the upgoing region on the roughly 5% level. Only in the downgoing region is significant uncertainty of about 14% observed, which can be attributed to the muon background. With these results in place the full data sample is unblinded in the following section. Results are shown in fig. 8.47

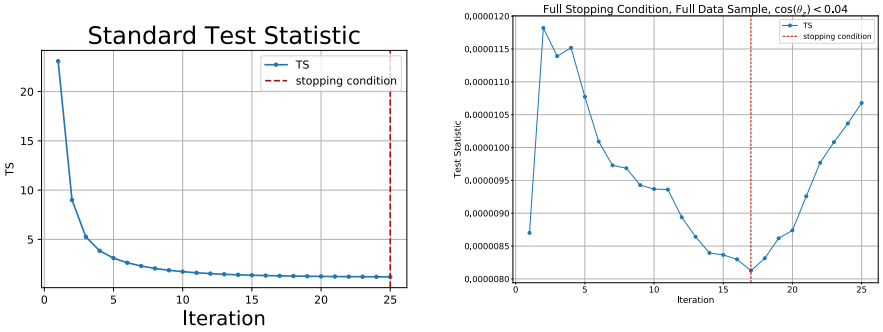


Figure 8.45: Stopping condition for the full data sample unfolding. Left: full stopping condition. The distribution follows scenario E in section 8.6 and the stopping condition comes to 17 iterations. Right: statistics only standard test statistic, which converges as expected on 25 iterations.

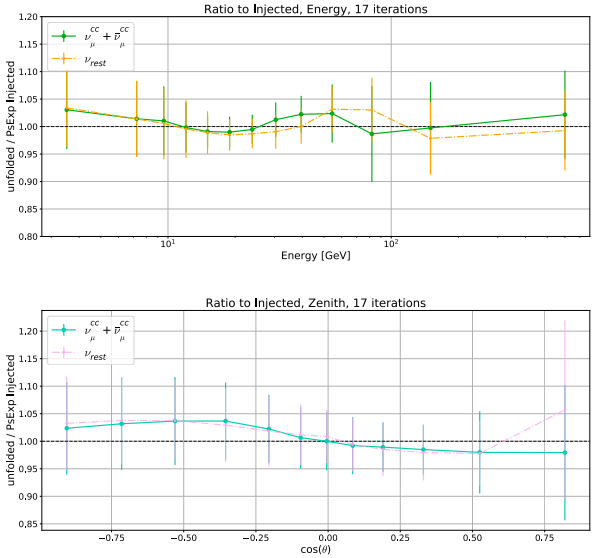


Figure 8.46: TSU test Ratio of unfolded to unfolded to injected pseudo experiment truth. Top: energy. Bottom: zenith angle. All ratios show agreement with the injected pseudo experiment within the standard 1σ error bars, with central values falling within 2.5% of pseudo experiment truth.

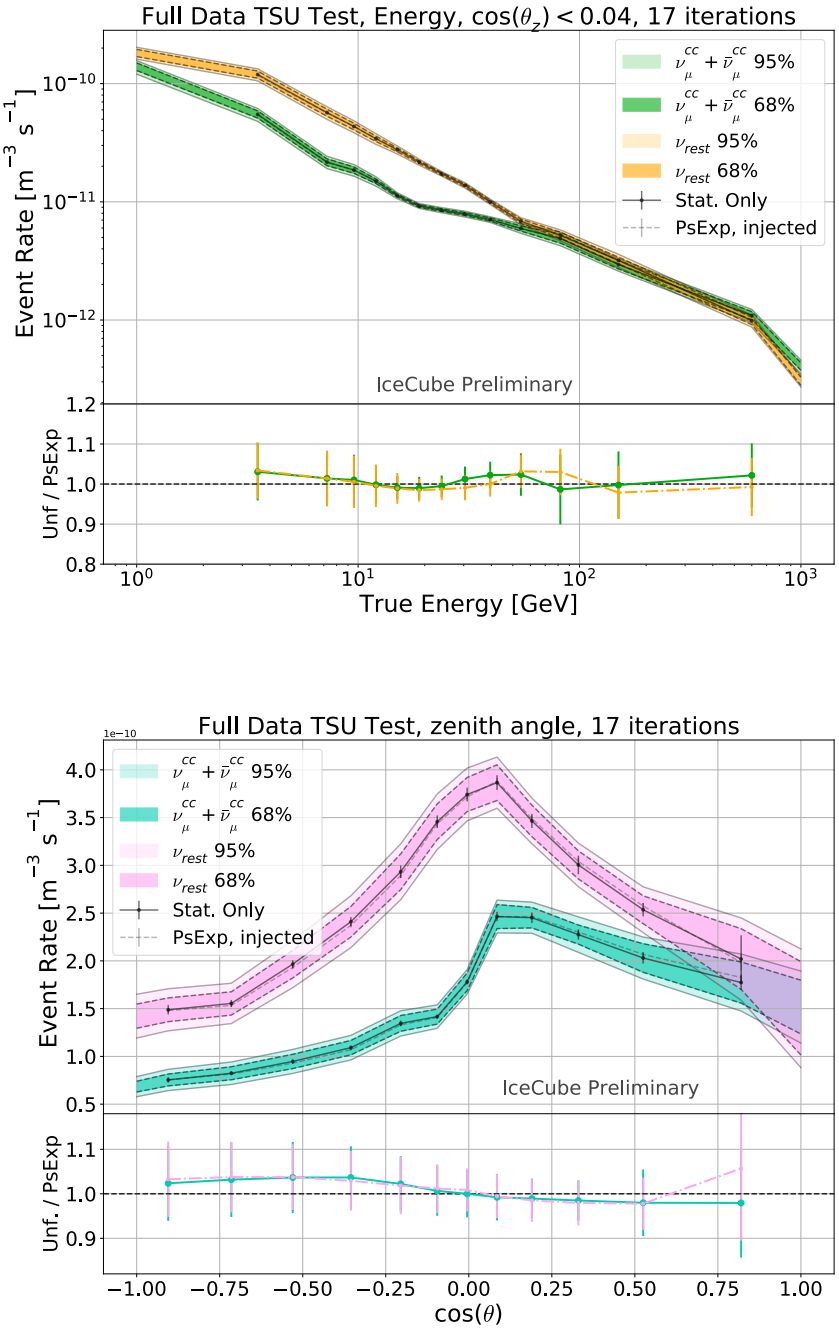


Figure 8.47: Result of the full data TSU test. Marginalized spectra for energy (left) and zenith angle (right) unfolding.

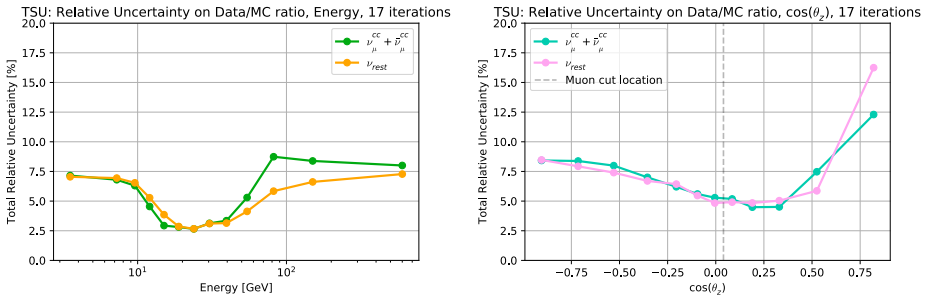


Figure 8.48: Relative uncertainty on the ratio of unfolded and injected pseudo experiment truth. Left: energy. Right: zenith angle.

8.16 Full Data Unblinding

With the above checks in place and the stopping condition set at 17 iterations, the full data sample is unblinded. A total of 204847 events are included in the data sample and are unfolded. The unfolded results are shown in fig. 8.50 along with the prediction from the HKKM model [7]. The relative uncertainties are shown in fig. 8.49. These results and their implications will be discussed in the following chapter.

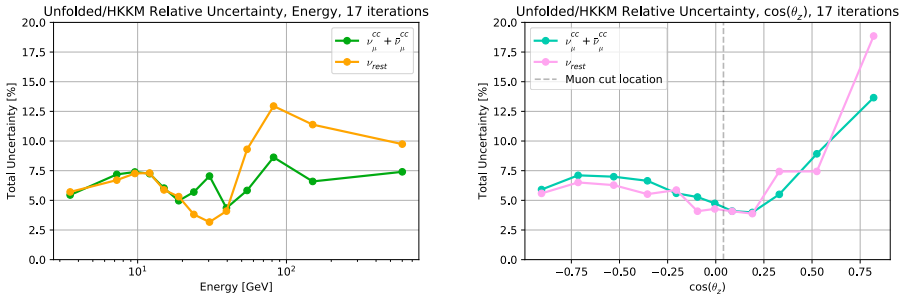


Figure 8.49: Relative uncertainty on the ratio of the unblinded unfolding and the standard prediction based on the HKKM model. Left: energy. Right: zenith angle.

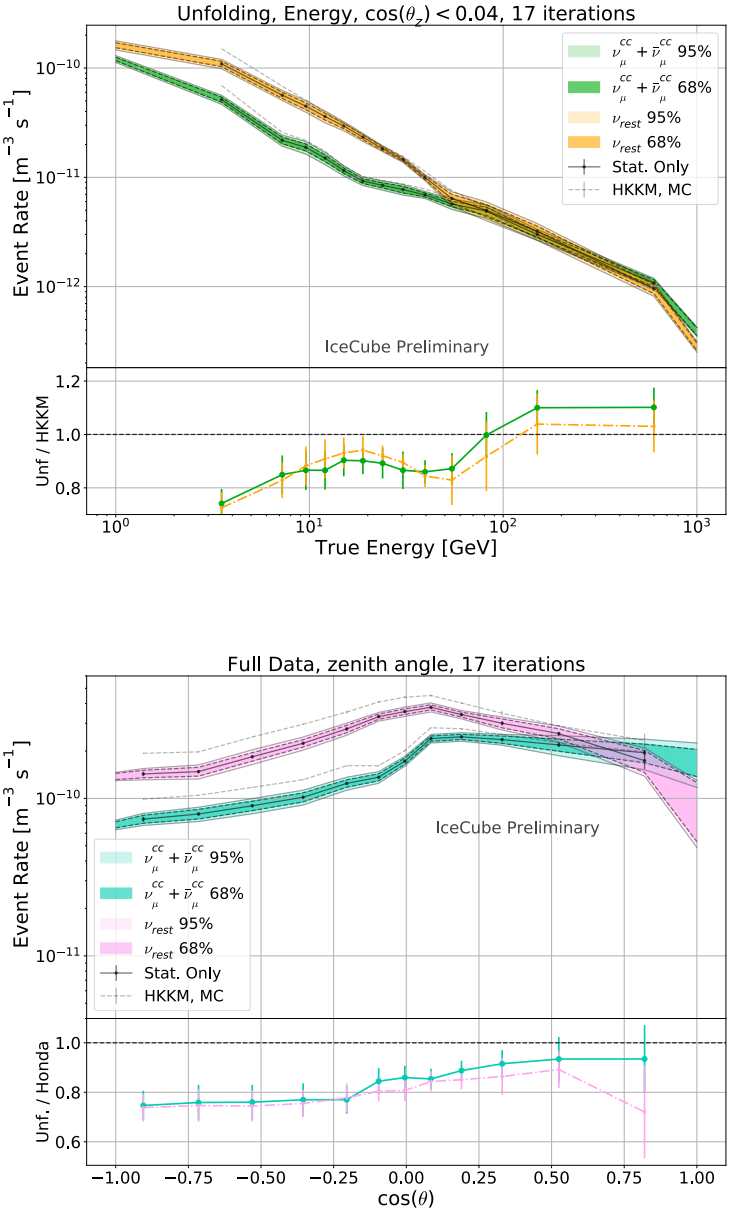


Figure 8.50: Full data sample unfolded results. The ratio is taken with respect to the HKKM simulation [7] which the unfolding matrix was based upon. Upper: energy. Bottom: zenith angle.

9 Discussion and Outlook

Presented in this thesis are the results of an iterative unfolding analysis, using 5 years of IceCube/DeepCore data, unfolding the atmospheric neutrino event rate by volume. Rigorous testing of the analysis method has been performed and after unblinding 204847 neutrino candidate events are included in the data sample. The unfolding method has been shown to be consistent within the standard 1σ error bars through a series of closure tests as described in chap. 8. This chapter will focus on the results, their implication and future outlook.

9.1 The Unfolded Result

The full unblinded and unfolded data set is shown in fig. 8.50, with the ratio relative to the MC prediction shown in the lower panel. The relative uncertainty on the unfolded/MC ratio is shown in fig. 8.49 and ranges from 3% to about 9%, except for the downgoing region where it reaches up to 18% for the cascade-like channel. However, in the region of interest to many oscillation experiments, below ~80GeV, the relative uncertainty on the ratio is very close to 3%. The relative uncertainty on the unfolded data itself also goes down to 3%, leading the way for good precision forward fits to this data in the future.

While displaying the kind of features which are expected from the neutrino flux, such as a power-law dependence, the unfolded data also show some tension in regards to the standard HKKM MC model. This is particularly evident in the energy range below 100 GeV and in the upgoing region, which is the region of interest when it comes to neutrino oscillations and matter effects. Where the MC model assumes values for the flux parameters, the oscillation parameters and the cross sections, the unfolding only compensates for the IceCube/DeepCore selection efficiency, volume and ice density. The discrepancy is most pronounced at energies below $\sim$ 60GeV and in the upgoing and horizon regions, which coincides with region of interest to neutrino oscillations. This could indicate an offset in the assumptions for the atmospheric neutrino oscillations and the matter effects, here referring to the various oscillation parameters such as Δm^2 and the electron density, when calculating the oscillations for the HKKM model. While the HKKM flux model was used to build the MC, it was decided the systematic uncertainties on the flux should not be included, as this would introduce systematic assumptions on the quantity of measurement. It is also important to note that when using the TSU test to verify the stopping condition, it has been shown that the choice of MC does not have a significant impact on the unfolded result, due to the nature of the iterative procedure. With that said and with the unfolding completed as shown in fig. 8.50, it would be illustrative to include the systematic uncertainty on the MC prediction. However, the systematic uncertainty is not included in the standard data set published by the HKKM authors. Only tentative estimates have been published in literature and presentations by the HKKM authors, indicating the systematic uncertainty to be on the scale of about 10% [9, 7]. An uncertainty of that scale is not large enough resolve the tension observed between the unfolded data and the HKKM model prediction. That said, there are investigations that can be performed in this dissertation in relation to the discrepancy, for example a direct comparison with the Super-K measurements.

9.2 Super-K Comparison

Super Kamiokande published an unfolding of the atmospheric neutrino flux in 2015 [3], using methods somewhat similar to those presented in this thesis. Important differences occur in the methodology used for setting the number of iterations, and the choice of unfolding channels. Where Super-K unfolds

into pure ν_μ and ν_e channels, this analysis unfolds into ν_μ^{cc} and ν_{rest}. The difference arises mainly from the exceptional muon tagging in Super Kamiokande, where the detected Cherenkov ring fuzziness is a strong discriminator between ν_μ and ν_e events. On the other hand IceCube/DeepCore can do rudimentary particle identification based on reconstructed track length, which does not have the same discriminative power. In order to keep the assumptions and statistical arguments in this work to a minimum, the ν_μ^{cc} and ν_{rest} channels were selected, as described in section 8.1. Although the unfolding channels differ a-priori between this work and the Super-K measurement, a direct comparison of the unfolding to MC agreement can none the less be meaningful, since the Super-K flux prediction only makes use of the measured and unfolded ν_μ^{cc} events, therefore the Super-k ν_μ channel is composed mostly of ν_μ^{cc} events. The ratio of the unfolded channels to the corresponding MC predictions are shown for the ν_μ^{cc} channel for Super-K [3] and for this analysis in fig. 9.1. Except for two outlying data points the results are statistically compatible within the standard 1σ error bars. A comparison between the Super-Kamiokande ν_e results and the unfolded ν_{rest} channel from this analysis is not valid since the latter also contains a ν_τ contribution, making the two particle selection channels different from each other. A calculation based on sample content by cross section fractions, was investigated but inflation of uncertainties quickly revealed this approach to be unfeasible. Notice however, that while the Super-K results are consistent with the HKKM prediction, there is some tension between the IceCube results and the HKKM prediction. A direct comparison of event rates would be more scientifically rigorous and potentially be able to reveal the scale of the latitude effect, however such data was not available from Super-K at the time of writing.

Super-K performs a basic closure test on the unfolding method using different inputs to set the number of iterations, and the work presented in this thesis expands heavily on the closure test methodology to achieve a more model independent method. Super-K presents their results in terms of the atmospheric neutrino flux, and their publication [3] gives the method for finding the flux from the unfolding as:

$$\frac{\Phi_{\text{unf}}}{\Phi_{\text{MC}}} = \frac{N_{\text{unf}}^{\text{cc}}}{N_{\text{MC}}^{\text{cc}}} \tag{9.1}$$

$$\Phi_{\text{unf}} = \frac{N_{\text{unf}}^{\text{cc}}}{N_{\text{MC}}^{\text{cc}}} \Phi_{\text{MC}}, \tag{9.2}$$

where N represents the unfolded event rate. Note that in this definition from Super-K only the charged current channel is used, allowing for one to one comparison with the ν_μ^{cc} result from this dissertation. The above relation however, relies on a few assumptions regarding the systematic parameters involved. The systematic parameters are assumed to completely cancel out both for the unfolding and the MC. While this is true in eq. 9.1, it is not in eq. 9.2. The strength of an unfolding analysis is to represent measurements in terms of 'true' data, but this estimate of the unfolded flux relies on the assumption that the systematic parameter values of Φ_{MC} are equal to those in nature, and the 'unfolded' flux estimate then directly relies on the assumptions of the input simulation. Even so, this is a simple and convenient estimate to present based on the measured ratio to MC and for comparison to the Super-K results, a similar calculation is included here, the result of which is shown in fig. 9.2. While the behavior of the unfolded data naturally follows the shape of the flux due to the calculation method, this plot is never the less illustrative of the unfolded data and the precision in the measurement, as the error bars are small on the y-axis scale (the relative uncertainty is shown in fig. 8.49). Similar illustrations have been produced by Super Kamiokande and a direct comparison of the ν_μ channel is presented in fig. 9.3. This again reveals the consistency between the Super-K and

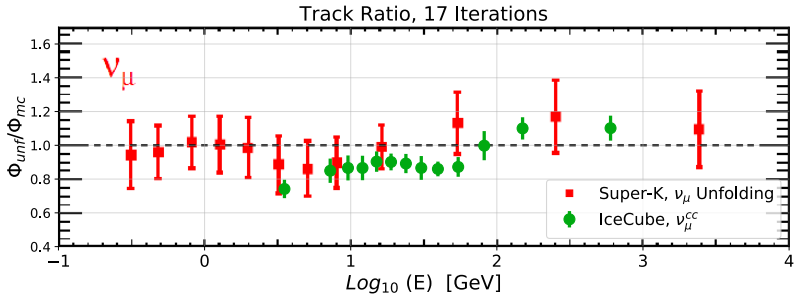


Figure 9.1: Data/MC agreement comparison between the unfolded ν_μ^{cc} channel and the corresponding result from Super Kamiokande result published in [3]. Note that the flux prediction is location dependent, and yields different fluxes for each site at energies below about 30GeV due to the latitude effect. This figure illustrates the agreement between the unfolded data and the model prediction at the South Pole and Kamioka respectively.

IceCube measurements, while highlighting the precision of the data presented in this work.

9.3 Further Developments

While the analysis presented herein is systematics dominated, the unfolding itself is only as good as the knowledge of the response matrix. More detailed knowledge of central systematics such as the DOM efficiency and ice properties are likely to have a significant impact on the results of analyses such as the one presented herein. Efforts are underway in the IceCube collaboration to address exactly these two systematics. Furthermore, building a larger Monte Carlo simulation and refining the sample selection with more accurate reconstruction algorithms would greatly benefit the outcome of the analysis. A sample selection living up to those requirements, named OscNext, have been in preparation alongside this analysis. While a few adjustments would have to be made to the analysis software used, applying the same analysis to the OscNext sample would very likely yield a reduction in the uncertainties shown in fig. 8.48.

9.3.1 Forward Folding

One way of investigating the discrepancy between the unfolded data and the HKKM model prediction would be to make a classic forward folding analysis and fit for the relevant parameters, such as the flux spectral index and the oscillation parameters, using a standard negative log likelihood approach on this unfolded data set. Some authors suggest avoiding forward fitting to an unfolded data, mostly In such an analysis a proper and detailed treatment of the systematic uncertainties would be of paramount importance. The scale of such an endeavor is beyond the scope of this work, and is suitable for a complete thesis on its own.

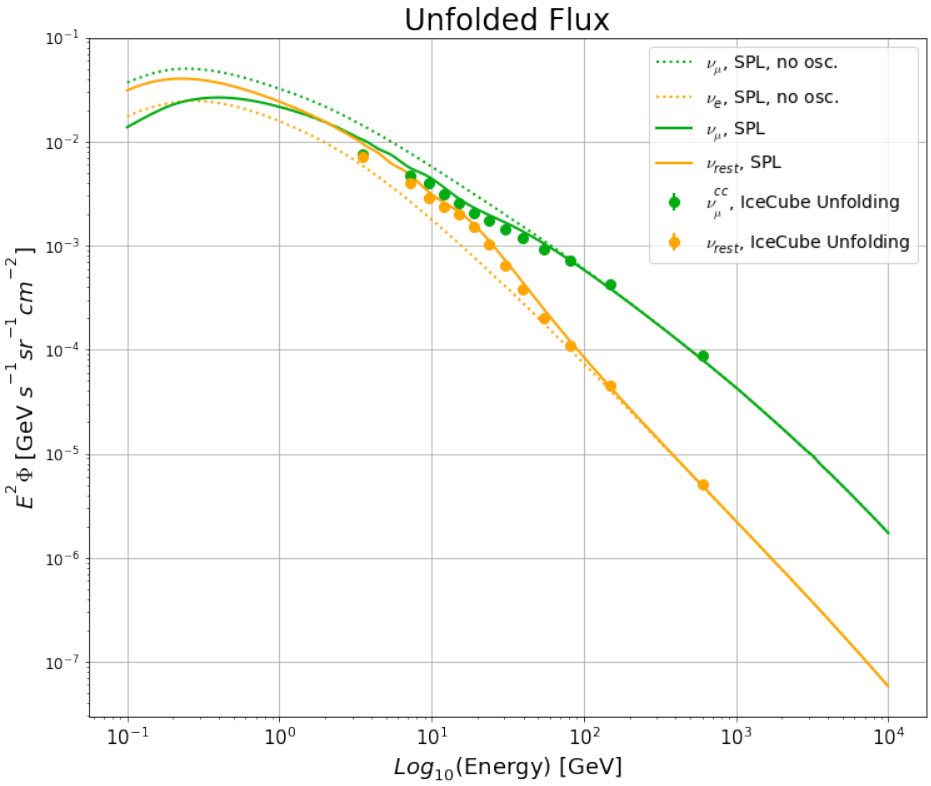


Figure 9.2: Illustrating the unfolded results in terms of atmospheric neutrino flux. Error bars are included on the data points.

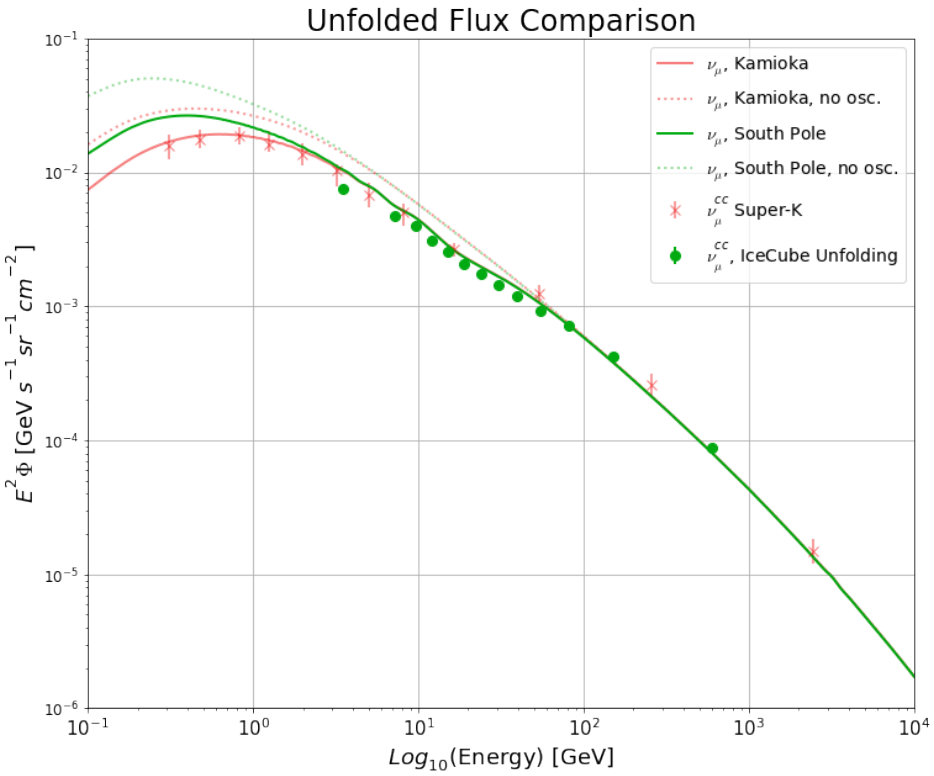


Figure 9.3: Comparison of the unfolded ν_μ flux between this work and Super Kamiokande. Data taken from [3] using webplotdigitizer [75]. Error bars are included on the data points for both the Super-K and IceCube unfolding, however the the IceCube error bars are tiny.

9.3.2 Direct-to-Flux Unfolding

It is in principle possible to unfold directly to the flux by modifying the truth side of the response matrix in order to compensate for the neutrino cross section by flavor and interaction type. In practice this is complicated by the following: the cross sections in the energy range of interest are not known to a high precision, which results in a large level of uncertainty. Preliminary tests performed during this work have indicated the impact of systematic uncertainty on the cross sections to be large enough that the unfolding diverges. This in turn leads the TSU test to set the number of iteration steps to be very low, at either one or two iterations, where the unfolding procedure is heavily biased towards the assumed MC. There is also the issue of a consistent systematics treatment: with the cross section included, several new systematic parameters would need to be included in the analysis as well. Using the random trial method would require a geometric increase in the number of trials to the point that it would become a significant challenge regarding computation time. Ideally the analysis presented in this work, having 4 major contributing systematics, would require on the order of 10^4 systematic trials to give a realistic picture of the systematics impact. With, for example, four new parameters this rises to 10^8 trials making the random sampling approach unfeasible. Instead a bracketing approach could be implemented (this was discussed and investigated for the work in this thesis as well), however, cross correlations between systematics and non-linear unfolding response to certain systematics complicates such an approach and can very easily lead to gross mis-estimates of the systematic uncertainty. Based on the above investigations and considerations, a decision was made for the work in this thesis to focus solely on unfolding the event rate. A direct to flux unfolding utilizing the methods presented in this thesis may later become feasible, provided sufficient progress is made in terms of computing power and neutrino interaction cross section knowledge.

9.4 Conclusions

The unfolding measurement of the atmospheric neutrino flux presented in this work contributes to the world knowledge of the atmospheric flux, and an updated overview of flux measurements is shown in fig. 9.4. Particularly, this work measures in a region of the energy spectrum that is otherwise sparsely populated, thereby enhancing the data resolution in said region. With regard to the objectives set out in chapter 1, the three main goals have been fulfilled: the uncertainty on the measurement, including systematic uncertainties, was brought down to between 3% and 7% in the oscillation region, as shown in fig. 8.49. This is a marked improvement over the ~20% uncertainty of the previous low energy measurement performed by Super Kamiokande [3]. Secondly, the 12 data points from this measurement adequately fill out the sparsely populated region between 10 GeV and ~100 GeV. Thirdly, the unfolding procedure was developed to provide an as model independent as possible measurement. Aside from these points, the spectrum was also unfolded in $\cos(\theta_z)$, with an uncertainty between 3% and 7% in the upgoing region, while the downgoing region had uncertainties of up to 18% due to muon contamination. The low dimensionality usually employed in a Bayesian unfolding is a point of critique within the Atlas collaboration, and something this analysis improves upon by unfolding in three dimensions [76]. Thus, this measurement provides a two dimensional unfolding for both particle identification channels in energy and $\cos(\theta_z)$, which will be of great interest for model builders when developing multi dimensional models of the atmospheric neutrino flux. Furthermore, this work goes beyond previous efforts in finding the optimum stopping condition for the number of iterations, which is sometimes selected arbitrarily in literature [77]. It could be argued that closure tests such as the TSU test presented in this work should be adopted as the industry standard when performing iterative Bayesian unfolding. In closing, it is remarked that this thesis presents the most precise

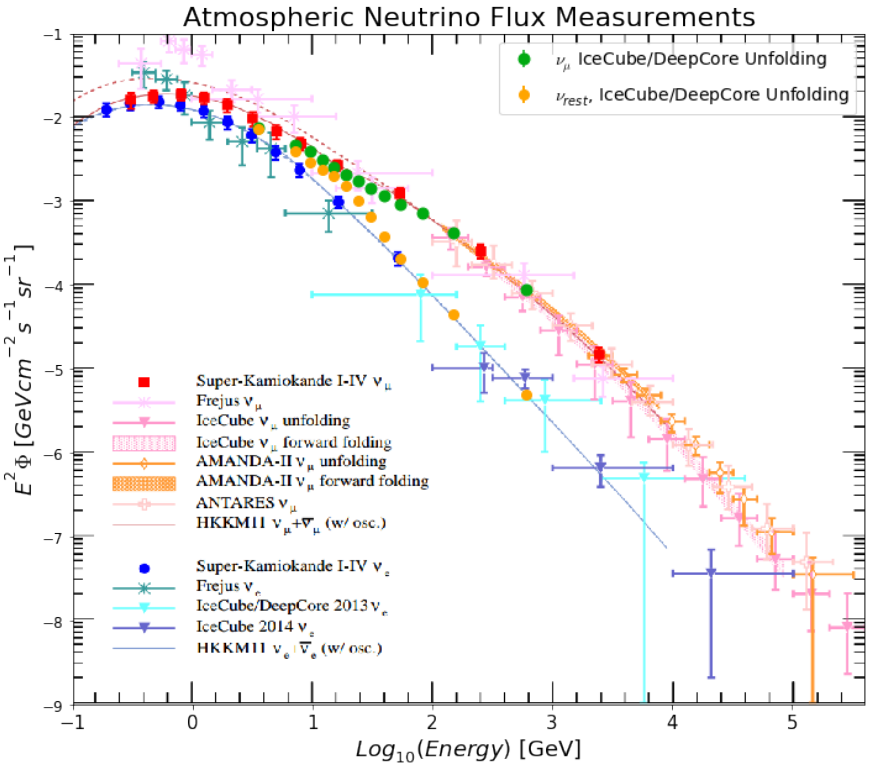


Figure 9.4: Updated overview of neutrino flux measurements, with the results from this analysis shown in green and orange points. The deviation of the ν_{rest} channel shown in orange from the standard ν_e prediction shown in blue in the energy region from $10-50$GeV is due to the cosmic ray latitude effect and the inclusion of ν_τ in the ν_{rest} channel. The expected fluxes at the South Pole and Kamioka are shown in fig. 1.4

unfolding measurement of the atmospheric flux at the time of publication, improves knowledge of the atmospheric flux in the region of interest to oscillation measurements, and has provided an as model independent as possible data set for future use by model builders.

Curriculum Vitae, Joakim Sandroos

Personal Information:

Name:	Joakim Sandroos
Date of Birth:	08th of May, 1980
Address:	Klitmøllervej 29 st.th, 2720 Vanløse
Country:	Denmark
Nationality:	Danish
Telephone:	+4542327667
Email:	j.sandroos@gmail.com

Profile:

Data scientist with flair for model testing, numerical simulation and data analysis. Has solid knowledge of neutrinos, particle physics and statistical models. Skilled at software development, particularly with Python and Git. Experience with management of larger teams, focusing on a collaborative work environment.

Scientific Appointments:

2015 - 2019: **Scientific Staff** **IceCube Group, J. G. U. Mainz**
Developed analysis of IceCube DeepCore data in order to measure the atmospheric neutrino flux. Using an interactive unfolding technique, showed it was possible to do the measurement in a model independent way and get clean data samples for each neutrino flavor and interaction. World leading results presented at the International Cosmic Ray Conference 2019. Work also included software development and teaching of advanced subjects such as particle physics and cosmology.

2013 - 2014: **Research Assistant** **Niels Bohr Institute, University of Copenhagen**
Developed a fast particle identification method based on the LEP-era Thrust variables and implemented them in the IceCube software environment. Using vector-calculus a fast characterization of events in IceCube was performed based on the spatial and time information of which optical modules are hit.

Outreach and Teaching

I have been involved in several outreach activities, where I have played an active role in designing and executing illustrative hands-on experiences for the public. More recently I was involved with the "Science Pavilion" held at the Roskilde music festival in collaboration with CERN. I am a native Danish speaker and can engage with local students and participate actively in outreach projects involving the local population. Regarding teaching experience, I have designed my own course on special relativity which ran over a semester to critical acclaim, as well as taught more advanced subjects at J.G.U. Mainz, including particle physics and cosmology.

Appointments:

2013: **Associate Lecturer** **Danish University Extension, Copenhagen**
Designed and hosted a course on special relativity over one semester. Duties consisted of creating a syllabus, organizing the structure of the lectures and giving them, as well as going through thought experiments and derivations with the students.

2012 - 2013: **Associate Lecturer** **Niels Steensens Catholic High School, Copenhagen**
Teaching math and physics to high school students, one of the main challenges was motivating the students. This I overcame by utilizing creative examples and engaging experiments, sparking their interest. Tasks included hosting classes, creating a syllabus, designing tests and grading homework.

2006 - 2007: **Exhibition Guide** **Tycho Brahe Planetarium, Copenhagen**
Speaking to the curiosity of visitors and guiding them through the main installation pieces of the planetarium exhibition. The main task consisted of communicating science in a way which made complicated topics accessible to the general public.

Education

I have had a strong interest in astronomy, computer science and physics since an early age. During my studies I focused on a combination of particle physics, astronomy and cosmology. This gave me to a solid background in particle physics and I graduated with a MSc in the field.

Taking the Initiative

My time at KamLAND came about as I knew I wanted to work with neutrinos, but since we had no neutrino group at the Niels Bohr institute where I was studying at the time, I had to take some initiative to make that happen: While professor Mogens Dam, was not an expert in neutrino physics, he agreed to act as co-supervisor for the project and the upcoming Master's Thesis. I contacted the exchange division at Tohoku University in Japan, where I re-established an old exchange agreement which had been inactive for years. Tohoku U. Professor and director of the Research Center for Neutrino Science institute Inoue Kunio accepted me as a special guest student, and I was able to start my work in the KamLAND collaboration. These initiatives meant I essentially brought the topic of neutrino physics to the Niels Bohr Institute, which today has a thriving community surrounding the field.

Education Overview:

2014: **Neutrinos Underground and in the Heavens** **NBI, University of Copenhagen**
Summer school on neutrino physics, organized by professor Subir Sarkar.

2013: **Coursera Course** **IPMU, University of Tokyo**
Completed the course:"From The Big Bang to Dark Energy", on particle physics and cosmology. Passed with distinction.

2012: **M.Sc degree in Physics** **Niels Bohr Institute, University of Copenhagen**

Master Thesis: "Two-Flavour Neutrino Oscillations and Geo-Neutrino Detection". The study was performed using data from the KamLAND experiment, and included working with large data samples, data analysis, background estimation and removal as well as development and testing of statistical methods. Main tools used were the ROOT analysis toolkit and C++. Course work pertaining to the degree has focused on particle physics and cosmology. Grade average: 11.2 (danish scale)

2009: **B.Sc degree in Physics** **Niels Bohr Institute, University of Copenhagen**

Bachelor's Thesis: "Method for Detection of Galaxy Cluster Shapes: A Study of A1689". Work was done using data from the CHADRA X-ray observatory, as well as simulated spectra from the XSPEC package. Work included numerical simulations, as well as closure tests and statistical methods.

2008: **JYPE** **Tohoku University, Sendai**

Completed the "Junior Year Program in English (JYPE)", at Tohoku University in Japan. Underwent individual research training at the KamLAND experiment, and took part in both data analysis and on-site detector upgrades. Also completed studies on Japanese language and culture.

Project Management

I have held several positions where organizational skills were important in regards to managing both teams and deadlines. Most notably I have been the CEO of an organization hosting an annual convention of about 3500 participants, alongside my other activities. This required managing several working groups in order for them to come together as a whole. Here the focus was on a positive work environment which served to effectively motivate our volunteers.

2010 - 2017: **CEO, head of finance, board member** **J-Popcon**

Spearheaded the Japanese culture convention J-Popcon, held annually in Copenhagen, and hosting about 3700 guests. Since 2013 part of the main board of organizers , where I took a leading role on new initiatives, logistics and finance. 2015 and 2016 appointed Chief Executive Officer, carrying the main responsibility for organizing the event, managing teams and making sure deadlines were met.

2008 - 2011: **Board Member** **Astronomy Student Association, University of Copenhagen**

Organizing astronomy related events for students at the University of Copenhagen.

Personal Skills

Language Skills:

Danish	Native speaker
English	Speaking, reading, writing: Native Level
Swedish	Speaking, reading, writing: Conversational Level
Norwegian	Speaking, reading, writing: Conversational Level
German	Speaking, reading, writing: Daily Conversational Level
Japanese	Speaking, reading: Certified to Level N4
.	

Software development

As part of an early project at NBI I implemented a software algorithm for fast particle identification in IceCube, based on the LEP-era 'Thrust' variables. This gave me understanding of IceCube software and I am familiar with developing my own IceTray modules. I have experience working in larger teams as well, as I have been part of the PISA framework development team in the early stages. This meant coming up with design choices on a larger scale and being part of implementing them via Github.

Programming & IT:

- Experienced user of **Python** - used for analysis, scripts and data analysis.
- Experienced user of **C++** in a scientific context. Used for analysis and simulations.
- Familiar with **Github** and **SVN** for version control.
- Experience working with big data processing, using supercomputer clusters, **condor**.
- Experience with **SQL**
- Familiar with **Perl** for parallelizing processes.
- Experienced user of **Bash**.

Software:

- Analysis Environments: Professional user of **Python** and **Pyplot**, the CERN **ROOT** framework used in particle data analysis as well as **MatLab**. Familiar with the **FITS** and **XSPEC** packages for generating and analyzing astrophysical spectra.
- Operating Systems: Experienced user of **Linux** (Ubuntu, Gentoo, Android) and **Windows** (Various versions)
- Visual Design: Familiar with Adobe **Photoshop**, Adobe **Illustrator**, **Inkscape** and **Sony Vegas**, used for producing graphic illustrations.
- Productivity Suite: Experienced user of **Ms Office**, **Libre Office**, **LaTeX**, **Overleaf** and **Lyx**

References

[1] Newton, Isaac. "Letter from Sir Isaac Newton to Robert Hooke". Historical Society of Pennsylvania. Retrieved 7 June 2018. https://discover.hsp.org/Record/dc-9792/Description#tabnav

[2] Daum, K. et al., Fréjus Collaboration, "Determination of the atmospheric neutrino spectra with the Fréjus detector", https://doi.org/10.1007/BF01556368

[3] E. Richard et al., Super Kamiokande Collaboration, "Measurements of the atmospheric neutrino flux by Super-Kamiokande: energy spectra, geomagnetic effects, and solar modulation", Phys. Rev. D 94, 052001 (2016), DOI: 10.1103/PhysRevD.94.052001

[4] Victor F. Hess, "Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten", Phys.Z. 13 (1912) 1084-1091

[5] Victor Hess, Alessandro De Angelis, "On the Observations of the Penetrating Radiation during Seven Balloon Flights",

[6] M. Tanabashi et al. (Particle Data Group), "2019 Review of Particle Physics", Phys. Rev. D 98, 030001 (2018) and 2019 update. https://journals.aps.org/prd/abstract/10.1103/PhysRevD.86.010001

[7] M. Honda et al. "Atmospheric neutrino flux calculation using the NRLMSISE-00 atmospheric model", Phys. Rev. D92, 023004 (2015), https://arxiv.org/abs/1502.03916

[8] A. Fedynitch, MCEq - Matrix cascade equations, https://github.com/afedynitch/MCEq

[9] M. Honda, "HKKM atmospheric neutrino fluxes", 1st Atmospheric Neutrino Workshop (ANW'16), https://indico.ph.tum.de/event/3533/contributions/2197/

[10] M. Honda, T. Gaisser, "Flux of Atmospheric Neutrinos", Ann.Rev.Nucl.Part.Sci.52:153-199,2002, https://arxiv.org/abs/hep-ph/0203272

[11] Andreas Höcker and Vakhtang Kartvelishvili, "SVD Approach to Data Unfolding", arXiv: 9509307v2

[12] Francis Halzen and Alan D. Martin, "Quarks and Leptons: An Introductory Course in Modern Particle Physics", Wiley; 1st edition (January 20, 1984), ISBN-10 : 0471887412

[13] David J. Griffiths, "Introduction to Elementary Particles", Second, Revised Edition, Wiley-VCH; 2nd edition (October 13, 2008), ISBN-10 : 9783527406012

[14] "Particle Data Group", official website https://pdg.lbl.gov/index.html

[15] Wikimedia Commons, Wikipedia, https://en.wikipedia.org/wiki/Standard_Model#/media/File:Standard_Model_o

[16] Ralph Engel, "Very High Energy Cosmic Rays and Their Interactions", Summary talk given at XIII ISVHECRI, Pylos (Greece), Sept. 2004; 22 pages, 8 figures, to appear in Nucl. Phys. B (Proc. Suppl.), DOI: 10.1016/j.nuclphysbps.2005.07.079

[17] A. H. Compton and R. N. Turner, "Cosmic Ray on the Pacific Ocean", Phys. Rev. 52, 799 – Published 15 October 1937, https://doi.org/10.1103/PhysRev.59.11

[18] Thomas H. Johnson, "The East-West Asymmetry of the Cosmic Radiation in High Latitudes and the Excess of Positive Mesotrons", Phys. Rev. 59, 11 – Published 1 January 1941, https://doi.org/10.1103/PhysRev.59.11

[19] Frank G. Schröder, "News from Cosmic Ray Air Showers", ICRC conference proceedings, Proceedings of Science, 358, 2019, https://pos.sissa.it/358/030/

[20] R. L. Sears,"Helium Content and Neutrino Fluxes in Solar Models", Astrophysical Journal, vol. 140, p.477, 1964. http://articles.adsabs.harvard.edu/pdf/1964ApJ...140..477S

[21] J. N. Bachall, "Solar Neutrino Cross Sections and Nuclear Beta Decay", Physical Review, vol. 135, num. 1B, 1964. https://authors.library.caltech.edu/33286/1/BAHpr64d.pdf

[22] R. Davis jr, "An Attempt to Observe the Capture of Reactor Neutrinos in Chlorine-37R". Proc. 1st UNESCO Conf., I, 728, 1957

[23] B. T. Cleveland et al. "Measurement of the Solar Neutrino Flux with the Homestake Chlorine Detector", The Astrophysical Journal, 496 : 505È526, 1998 March 2, https://iopscience.iop.org/article/10.1086/305343/pdf

[24] GALLEX collaboration, "GALLEX Solar Neutrino Observations: Results for GALLEX IV", Physics Letters B Volume 447, Issues 1–2, 4 February 1999, Pages 127-133, https://doi.org/10.1016/S0370-2693(98)01579-2

[25] K. S. Hirata et al. Real-time, directional measurement of 8B solar neutrinos in the Kamiokande II detector.", Physical Review D 44, 2241-2260, 1991, https://doi.org/10.1103/PhysRevD.44.2241

[26] SAGE collaboration, "Measurement of the solar neutrino capture rate with gallium metal", Phys.Rev.C60:055801,1999, https://doi.org/10.1103/PhysRevC.60.055801

[27] Christian Iliadis, "Nuclear Physics of Stars", Weinheim : Wiley-VCH, ©2007.

[28] "proton-proton chain", wikipedia: https://en.wikipedia.org/wiki/Proton%E2%80%93proton_chain_reaction#/media

[29] K. Hirata et al. "Results from One Thousand Days of Real-Time, Directional Solar-Neutrino Data", Physical review letters 65, https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.65.1297

[30] A. Bellrive et al. "Sudbury Neutrino Observatory", Nuclear Physics B Volume 908, July 2016, Pages 30-51, arXiv:1602.02469 [nucl-ex]

[31] B. Pontecorvo, "Mesonium and antimesonium," Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki, vol. 33, p. 549, 1957, Soviet Physics—JETP, vol. 6, p. 429, 1958.

[32] B. Pontecorvo, "Inverse beta processes and nonconservation of lepton charge," Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki, vol. 34, p. 247, 1957, Soviet Physics—JETP, vol.7, pp. 172-173, 1958.

[33] M. Gell-Mann and A. Pais, "Behavior of neutral particles under charge conjugation," Physical Review, vol. 97, no. 5, pp. 1387–1389, 1955.

[34] T. Kajita, E. Kearns, M.Shiozawa, "Establishing Atmospheric Neutrino Oscillations with Super-Kamiokande". Nuclear Physics B 908, April 2016

[35] T. Gaisser, R. Engel, E. Resconi, "Cosmic Rays and Particle Physics", Cambridge University Press, 2016, ISBN 10: 0521016460

[36] F. Reines et al. "Evidence for High-Energy Cosmic-Ray Neutrino Interactions", Physical Review Letters D 15, 1965, p. 429-433. https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.15.429

[37] J. A. Formaggio, G. P. Zeller, "From eV to EeV: Neutrino Cross-Sections Across Energy Scales", Rev. Mod. Phys. 84, 1307 (2012), https://arxiv.org/pdf/1305.7513.pdf

[38] J. L. Hewett et al. "Fundamental Physics at the Intensity Frontier", ANL-HEP-TR-12-25, SLAC-R-991, https://arxiv.org/pdf/1205.2671.pdf

[39] C. V. Achar et al. "Detection of muons produced by cosmic ray neutrinos deep underground". Physics Letters, Volume 18, Issue 2, p. 196-199. https://doi.org/10.1016/0031-9163(65)90712-2

[40] M. Honda, T. Kajita, S. Midorikawaand K. Kasahara, "Calculation of the flux of atmospheric neutrinos", Phys. Rev.D52, 4985(1995).

[41] Y. Fukuda et al., "Atmospheric $\nu_\mu\nu_e$ ratio in the multi-GeV energy range", Physics Letters B Volume 335, Issue 2, 1 September 1994, Pages 237-245, https://doi.org/10.1016/0370-2693(94)91420-6

[42] Particle Data Group, "CKM Quark-Mixing Matrix", http://pdg.lbl.gov/2019/reviews/rpp2018-rev-ckm-matrix.pdf

[43] Prof. Dr. Dr. h.c. Manfred Lindner, "Astroteilchenphysik in Deutschland:Status und Perspektiven 2005"

[44] R. Wendell, Prob3++ software for computing three flavor neutrino oscillation probabilities (2012–). https://webhome.phy.duke.edu/∼raw22/public/Prob3++/

[45] Particle Data Group, "Neutrino Masses, Mixing, and Oscillations", http://pdg.lbl.gov/2019/reviews/rpp2018-rev-neutrino-mixing.pdf

[46] M. Dracos, "THE EUROPEAN SPALLATION SOURCE NEUTRINO SUPER BEAMDESIGN STUDY", http://accelconf.web.cern.ch/AccelConf/ipac2018/papers/tupml068.pdf

[47] A. Smirnov, "The MSW Effect and Matter Effects in Neutrino Oscillations", Phys.Scripta T121 (2005) 57-64, DOI: 10.1088/0031-8949/2005/T121/008 https://arxiv.org/pdf/hep-ph/0412391.pdf

[48] A. Donini et al. "Neutrino tomography of Earth", Nature Phys 15, 37–40 (2019), https://doi.org/10.1038/s41567-018-0319-1

[49] C. Giunti and C.W. Kim, "Fundamentals of Neutrino Physics and Astrophysics", Oxford University Press, 2007, ISBN: 9780198508717

[50] P.F. de Salas et al. "Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity", Physics Letters B, vol 787, July 2018, https://doi.org/10.1016/j.physletb.2018.06.019

[51] "M. Aartsen et al., "The IceCube Neutrino Observatory: Instrumentation and Online Systems", Journal of Instrumentation, arXiv:1612.05093 [astro-ph.IM](03), December 2016

[52] IceCube Collaboration, "The Design and Performance of IceCube DeepCore", Ap.J. 833, 3, Dec. 2016, https://arxiv.org/abs/1109.6096

[53] "M Aartsen et al., "Letter of Intent: The Precision IceCube Next Generation Upgrade (PINGU)", arXiv: 1401.2046 [physics.ins-det], 2014

[54] Jacob Daughhetee, "DeepCore Filter Proposal", IceCube/DeepCore internal communication, 2012, https://docushare.icecube.wisc.edu/dsweb/Get/Document-62751/filter_proposal3.pdf

[55] M. R. Grootendorst et al. "Cerenkov luminescence imaging (CLI) for image-guided cancer surgery", Clin Transl Imaging 4, 353–366 (2016), https://doi.org/10.1007/s40336-016-0183-x

[56] Hadiseh Alaeian, "An Introduction to Cherenkov Radiation", http://large.stanford.edu/courses/2014/ph241/alaeian2/, 2014

[57] C.Andreopoulos et al., "The GENIE neutrino Monte Carlo generator", https://doi.org/10.1016/j.nima.2009.12.009, https://github.com/GENIE-MC/Generator

[58] M. Leuermann. "Testing the Neutrino Mass Ordering with Three Years of IceCube/DeepCore Data". Pending. Ph.D. Rheinisch-Westfalische Technische Hochschule (RWTH) Aachen, 2018

[59] M. Dunkman. "Measurement of Atmospheric Muon Neutrino Disappearance with IceCube-DeepCore". Ph.D. The Pennsylvania State University, 2015

[60] Michael Larson, PhD Thesis

[61] Yann Coadou, "Boosted Decision Trees and Applications", EPJ Web of Conferences Volume 55, 2013 SOS 2012 – IN2P3 School of Statistics, https://doi.org/10.1051/epjconf/20135502004

[62] AMANDA Collaboration. "Muon track reconstruction and data selection techniques in AMANDA". Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 524.1 (2004), pp. 169 –194

[63] S. Euler. "Observations of Oscillations of Atmospheric Neutrinos with the IceCube Neutrino Observatory". Ph.D. Rheinisch-Westfalische Technische Hochschule (RWTH), 2013

[64] T.PierogaK.Wernerb, "EPOS Model and Ultra High Energy Cosmic Rays", Nuclear Physics B, Volume 196, December 2009, Pages 102-105, https://doi.org/10.1016/j.nuclphysbps.2009.09.017

[65] J. van Santen. "Neutrino Interactions in IceCube above 1 TeV: Constraints on Atmospheric Charmed-Meson Production and Investigation of the Astrophysical Neutrino Flux with 2 Years of IceCube Data taken 2010–2012". PhD thesis. U. Wisconsin, Madison (main), 2014

[66] D. Pandel. "Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal Neutrinoteleskop NT-72". PhD thesis. Humboldt-Universitat, 1996

[67] "M. Aartsen et al. "The IceCube Neutrino Observatory Part II: Atmospheric and Diffuse UHE Neutrino Searches of All Flavors", https://arxiv.org/pdf/1309.7003.pdf

[68] M. J. Larson. "Simulation and identification of non-Poissonian noise triggers in the IceCube neutrino detector". M.Sc. Alabama U., 2013

[69] Guenter Zech, "Analysis of distorted measurements - parameter estimation and unfolding", arXiv: 1607.06910v1

[70] G. D'Agostini, "Improved iterative Bayesian unfolding", arXiv: 1010.0632v1

[71] Shimazaki and Shinomoto. Neural Comput, 2007, 19(6), 1503-1527

[72] T. Adye, "Unfolding algorithms and tests using RooUnfold", arXiv: 1105.1160

[73] M. Ackermann et al., "Optical properties of deep glacial ice at the South Pole", D13203 Journal of Geophysical Research Atmospheres, Volume 111, Issue number 13, 2006

[74] IceCube Collaboration, "Measurement of South Pole ice transparency with the IceCube LED calibration system", https://arxiv.org/abs/1301.5361

[75] Ankit Rohatgi, "Web Plot Digitizer v.4.2", https://apps.automeris.io/wpd/

[76] Dr. Jørgen Bech (associate professor), Private Communication, Niels Bohr Institute, University of Copenhagen,

[77] Georgios Choudalakis, "Unfolding in ATLAS", April 2011, https://arxiv.org/pdf/1104.2962.pdf

[78] W.Feller, "An introduction to probability theory and its applications", Wiley; 3rd edition (1968), ISBN-13: 978-0471257080

[79] T. Ruhe, Private Communication