STUDIES OF THE ENERGY, AZIMUTHAL, AND TIME SPECTRA OF THE ATMOSPHERIC NEUTRINO FLUX AT SUPER-KAMIOKANDE

EUAN RICHARD

ABSTRACT

This thesis presents three direct measurements of the atmospheric $\nu_e +$ $\bar{\nu}_e$ and $\nu_\mu + \bar{\nu}_\mu$ neutrino fluxes at Super-Kamiokande: the directionally-integrated energy spectra, the azimuthal spectra, and the modulation of the fluxes with time over the 11-year solar cycle.

In particular, the energy spectra in the sub-GeV to TeV range was measured and compared to the predictions of various published flux models. While none of the models were strongly inconsistent with the data, some preference was seen for the HKKM11 model as the most realistic model. The neutrino and antineutrino fluxes were also examined separately, and found to be consistent with the models.

The azimuthal analysis measured the east-to-west asymmetry in the neutrino flux, caused by the geomagnetic field, for both flavours at $> 5\sigma$. Measurements were made of the strength of the effect as functions of energy and zenith angle. There was also an indication that the alignment of the asymmetry was dependent on zenith angle, seen at the 2.2 σ level.

A search for a long-term correlation between the atmospheric neutrino flux and the solar magnetic activity cycle was performed, however the expected effect based on the HKKM model was calculated to be relatively small. An indication of a correlation was seen at the 1.1 σ level. During particularly strong solar activity events, known as Forbush decreases, no theoretical prediction is available, but a deviation from the expected neutrino event rate is seen at the 2.4 σ level.

ACKNOWLEDGMENTS

So gladly we march to mines, tunnels or seas
New physics or not, we get PhDs –
Condensed to a preprint we'll mail to our mothers
Of one page of physics (two pages of authors).

— B. Nolty and K. Scholberg

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ATM   Analogue Timing Module

AFT   After Trigger

CC   Charged Current

CCQE Charged Current Quasi-Elastic

DIS   Deep Inelastic Scattering

DOF   Degrees of Freedom

FC   Fully Contained

FV   Fiducial Volume

GUT   Grand Unified Theory

HE   High Energy

HV   High Voltage

ICRR   Institute for Cosmic Ray Research

ID   Inner Detector

LE   Low Energy

MC   Monte Carlo

NC   Neutral Current

NM   Neutron Monitor

OD   Outer Detector

PC   Partially Contained

PDG   Particle Data Group

PDF   Probability Distribution Function

PID   Particle Identification

PMT   Photomultiplier Tube

QE   Quantum Efficiency

QE   Quasi-Elastic

RES   Resonant

SK   Super-Kamiokande

SM   Standard Model

SLE   Super Low Energy

SHE   Super High Energy

TQ   Time and Charge
TOF   Time of Flight
UPMU Upwards-going Muon
UTC   Coordinated Universtal Time

Part I

INTRODUCTION

I think, therefore I am.

— René Descartes

NEUTRINO PHYSICS

1.1 MOTIVATION: THE ATMOSPHERIC NEUTRINO

Atmospheric neutrinos are generated from the decay of mesons produced in cosmic-ray interactions in the Earth's atmosphere, and are one of the main experimentally available neutrino sources. Since the first detection of the atmospheric neutrino in underground experiments in the 1960s [1, 2], further measurements of them brought the discovery of neutrino oscillation (and thus finite neutrino masses) in 1998 [3]. Since then the understanding of the oscillation parameters has improved quickly, by using the atmospheric [4, 5] data in combination with solar [6, 7], reactor [8, 9, 10, 11], and accelerator [12, 13, 14, 15] sourced neutrinos.

The atmospheric neutrino sample spans a wide energy region, peaking at $O(\mathrm{MeV})$ but so far measured up to $O(\mathrm{PeV})$, and atmospheric neutrinos can be detected up to $\sim 13\,000\,\mathrm{km}$ away from their initial creation point (i.e. after crossing the diameter of the Earth). They are thus a useful source as a varied and high-statistics neutrino beam, however to understand and use them as a sample, an accurate prediction of the expected flux depending on neutrino flavour, energy, direction, and time is required. For example, the discovery of neutrino oscillation was by comparison of their expected zenith angle distribution with and without oscillation applied, as shown in Fig. 1 (although an up-down asymmetry can be predicted using some simple flux assumptions and geometrical considerations [16], a more detailed flux model is necessary for an accurate quantitative comparison). Furthermore, the atmospheric neutrino is not just a signal but also a background source to many other experiments, such as astrophysical neutrino, proton decay, and dark matter searches – one example is given in Fig. 2. An accurate prediction of the flux is thus of paramount importance to many current and future experiments. As one example, the planned Hyper-Kamiokande [17] experiment could use the atmospheric sample to uncover the neutrino mass hierarchy.

Current atmospheric neutrino flux predictions are given by Monte Carlo simulations by several groups [18, 19, 20, 21]. These models are based on data such as the primary cosmic ray proton flux and the secondary muon flux, measured in the atmosphere by particle detectors on balloons [22, 23] and spacecraft [24, 25], and hadron production measured in accelerator experiments [26, 27]. Direct experimental measurements of the flux energy spectra were made by the Frejus [28] collaboration (before neutrino oscillation was known), and
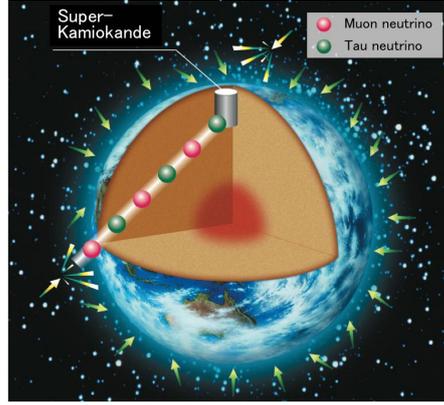
(a) Graphical illustration
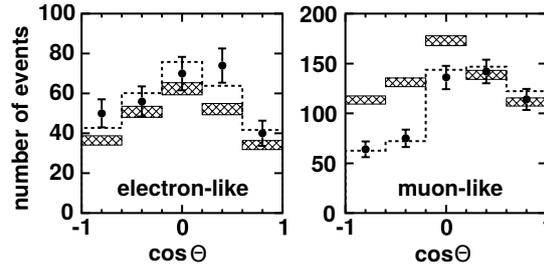
(b) Zenith distributions

Figure 1: The first strong evidence for neutrino oscillation found by the Super-Kamiokande detector. Zenith angle distributions show early data reported in 1998 for contained lepton events, where $\cos(\Theta) = 1$ represents down-going leptons. Dashed (hatched) values show the predictions based on neutrino flux models for an oscillated (unoscillated) flux. In particular, it appears that the up-going $\nu_\mu + \bar{\nu}_\mu$ flux has reduced on its way through the Earth, matching the model of oscillations to tau or sterile neutrinos.

more recently by the AMANDA-II [29, 30] and IceCube [31, 32, 33] collaborations at higher energies (up to the 100 TeV range).

The estimated uncertainties on the simulations, especially at lower energies, were historically much smaller than the uncertainties on the direct experimental measurements of the flux. However, as the current generation of neutrino detection experiments improve statistics and reduce systematic uncertainties, measuring the atmospheric neutrino flux by direct data measurements becomes increasingly useful – allowing more accurate cross-checks and feedback to the future development of the simulations, and discovery of several new effects that have been predicted by the simulations but not yet observed, such as directional asymmetries and modulation over time. Specifically this thesis presents three direct measurements of the atmospheric neutrino flux at Super-Kamiokande: the directionally-integrated energy spectrum of the flux (with high statistics in the 100 MeV to 10 TeV range), the azimuthal spectrum, and the modulation of the flux with time over the eleven-year solar cycle.
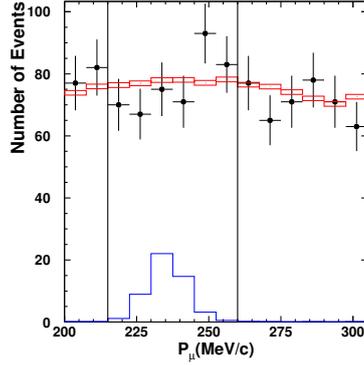
Figure 2: An example of the atmospheric neutrino flux as a background source – data from the search for proton decay via the mode $p \to \nu + K^+$ at Super-Kamiokande. The reconstructed muon momentum distribution is shown. The red boxes correspond to the atmospheric neutrino simulation, and the blue line histogram is the expected proton decay signal.

1.2 DETECTION OF THE NEUTRINO

1.2.1 *Theoretical basis*

In the Standard Model (SM), the neutrino is a spin-$\frac{1}{2}$ lepton which participates only in the gravitational and weak interactions. Depending on the neutrino's cosmological energy density, their gravitational effects may be observed on the large scale, e.g. as a minority component of dark matter [34]. Detection of individual neutrinos can be accomplished via the weak force, which is carried by the charged and neutral vector bosons $W^\pm$ and Z^0 with masses of 80 GeV and 91 GeV respectively, and thus has an effective range restricted to $\sim 10^{-17}$ m. Although the intrinsic strength of the weak force is much higher than the electromagnetic force, weak interactions are thus comparatively infrequent.

The possible interactions of a neutrino by the neutral current (NC) are simple ones of the form Fig. 3a, where the Z^0 boson facilitates a transfer of momentum and spin. Charged current (CC) interactions are more complicated, and involve changes in weak isospin (i.e. up-type vs. down-type quarks, or charged vs. neutral leptons) and flavour. The flavour changes originate due to the difference between propagation and interaction eigenstates; this will be considered in detail in Sec. 1.3. One such example is shown in Fig. 3b, where it can be observed that contrary to NC interactions, in CC interactions the flavour of the interacting neutrino is preserved in the outgoing charged lepton, and may be experimentally observable in a practical sense.

(a) Neutral current   (b) Charged current
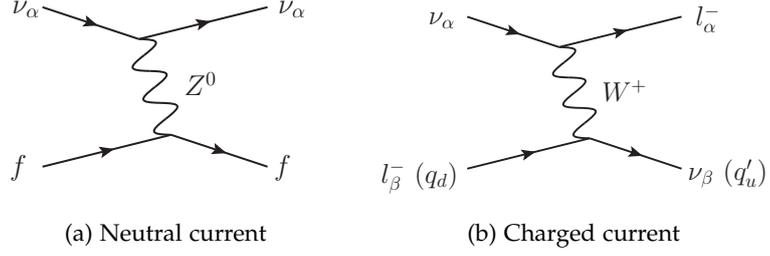
Figure 3: Examples of Standard Model interactions for a neutrino ν_α ($\alpha = e, \mu, \tau$). In (a), f represents any weakly-charged fermion such as a charged lepton, left-handed neutrino, or quark. In (b), l_β^- is a negatively charged lepton of flavour β, q_d represents a down-type quark, and q_u' an up-type quark after a flavour basis rotation (i.e. a superposition of u, c, and t).

Mathematically speaking, the propagator for the $W^\pm$ and Z^0 bosons is given by

$$\frac{-i(g_{ab} - q_\alpha q_\beta / m^2)}{q^2 - m^2} \tag{1}$$

where $g_{\alpha\beta}$ is the Minkowski metric, m is the mass of the $W^\pm$ or Z^0 boson, and q is the momentum associated to the boson. When $q^2 \ll m^2$, this reduces to $ig_{\alpha\beta}/m^2$, which is a pointlike interaction as postulated in Fermi's original theory of weak interactions. In other words, when the energy transfer is less than the mass of the $W^\pm$ and Z^0 bosons, we can ignore the effects of their propagation. The full matrix element is then

$$M_F = \frac{G_F}{\sqrt{2}} \left[\bar{\varphi} \gamma^\alpha (\mathbf{1} - \gamma^5) \varphi \right] \left[\bar{\varphi} \gamma_\alpha (\mathbf{1} - \gamma^5) \varphi \right] \tag{2}$$

where the definition of the overbar operator is $\bar{\varphi} = \varphi^\dagger \gamma^0$, G_F is the Fermi coupling constant, and $\frac{1}{2}(\mathbf{1} - \gamma^5)$ is the left-handed chiral projection operator. The use of the Fermi theory, although we now know as being simply an effective theory, is a useful simplification over a wide energy range. The presence of the operator $\frac{1}{2}(\mathbf{1} - \gamma^5)$ however bears close examination – it essentially represents the important experimentally-determined fact [35] that the weak interaction violates parity; only left-chirality fields (or right-chirality antiparticle fields) participate in it. This is often referred to as the *vector − axial vector* (or $V − A$) structure of the weak interaction, referring to the objects $\bar{\varphi}\gamma^\alpha\varphi$ and $\bar{\varphi}\gamma^\alpha\gamma^5\varphi$, which transform respectively as a four-vector and an axial vector under a Lorentz transformation (where an axial vector, or pseudovector, gains an additional factor of -1 under a parity flip).[1]

---

1 The other possible Lorentz-invariant currents are the scalar, pseudoscalar, and tensor currents, which are experimentally found to not be involved in the weak interaction.

From an experimental point of view, we observe neutrino interactions with hadronic targets such as water or iron, and a range of hadronic effects must be additionally considered; this is discussed in greater length in Sec. 4.5.1 when the Monte Carlo event generators are considered. Depending on the neutrino energy and target nuclei, in particular for neutrino energies of roughly 1 GeV, some CC reactions still take a simple form as in Fig. 3b and are termed "Quasi-Elastic" (CCQE) interactions. In this case, the energy transfer q^2 tends to be small, and the outgoing lepton can roughly preserve the momentum vector of the incoming neutrino; providing that the outgoing lepton can be well measured, we can thus reconstruct all of the neutrino's properties.[2]

1.2.2 *Experimental history*

The initial motivation for postulating the existence of the neutrino, by W. Pauli in 1930 [36], was to explain the "missing" momentum in the beta-decay process that we now know to proceed as $n \rightarrow p + e^- + \bar{\nu}_e$. The discovery, by Cowan and Reines in 1956 [37], used a similar but inverted reaction $\bar{\nu}_e + p \rightarrow e^+ + n$ which could take advantage of an intense $\bar{\nu}_e$ flux produced by a nuclear reactor. The actual detection method used was to monitor for a 2γ signal from positron annihilation, proceeded a few microseconds later by a γ signal emitted when the neutron is captured by a nucleus.

The discovery of a naturally-occurring neutrino flux came around a decade later, in the form of atmospheric neutrinos, by plastic scintillation detectors several kilometers underground [1, 2]. The depth of the experiments reduced the expected background of cosmic-ray secondary muons, allowing the signal events of neutrino-induced muons created in the surrounding rock to be more easily observed (although some care still had to be taken to avoid contamination by the remaining down-going background muons).

A few years later, the Homestake experiment [38] used a radiochemical detector in a mine to identify neutrinos produced in the Solar fusion process, which created a precisely-measurable number of Argon isotopes in the reaction $\nu_e + {}^{37}\text{Cl} \rightarrow {}^{37}\text{Ar} + e^-$. However, the flux was approximately a third of that expected by theoretical solar model calculations – due to the as-yet-undiscovered neutrino oscillations (modified by matter effects inside of the Sun). It was not until three decades later, with evidence from the water Cherenkov detectors Super-Kamiokande (SK) [3] and the Sudbury Neutrino Observatory (SNO) [39] that enough evidence was gathered to prove the exis-

---

2 Strictly speaking, even in the CCQE case the momentum of outgoing leptons can vary considerably compared to the neutrino's original momentum, but as the process is theoretically relatively simple to study much information can be extracted in a statistical sense.

tence of neutrino oscillations. These detectors observe the Cherenkov light emitted by charged leptons, which originate in neutrino collisions in large fiducial-volume water tanks (in the case of SNO, heavy water), and unlike previous detectors were able to accurately to reconstruct the momentum and flavour of individual neutrinos in near real-time (the SK case will be examined in detail in Sec. 4.1).

Several recent experiments have returned to reactor sourced neutrinos [8, 9, 10, 11], or constructed artificial neutrino beam generators [12, 13, 14, 15] in order to make precision measurements of the neutrino oscillation parameters.

The field of neutrino astronomy also received a kick-start with the detection of neutrinos originating from the supernova 1987A [40], located in the Large Magellanic Cloud, in agreement with models suggesting that 99% of a supernova's energy is released in the form of neutrinos. The discovery of a regular astrophysical neutrino flux has as of this year been confirmed by the gigaton-volume Cherenkov detector IceCube [41], which due to its huge volume is able to reconstruct events up to PeV energies.

1.3 NEUTRINO OSCILLATION

The oscillation of neutrino flavour during propagation is a comparatively recent and influential discovery, implying neutrino mass and several new parameters, and providing clear evidence for a necessary extension beyond the SM. A detailed description of the oscillation mechanism, parameters, and uncertainties is necessary for the measurements in this thesis, as we will be comparing the measured neutrino flux at the SK detector with the unoscillated predictions from the flux simulations.

1.3.1 *The lepton mixing matrix*

Neutrino oscillation measurements so far are consistent with the three-flavor lepton mixing scheme described by a unitary 3×3 PMNS matrix [42, 43]

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{3}$$

which represents a rotation in field space between the weak-flavour basis ν_α ($\alpha = e, \mu, \tau$) and the mass basis ν_i ($i = 1, 2, 3$), or in other words the interaction and propagation bases. Anti-particles can be described by the conjugate matrix. Crucially, if the masses m_i of each state ν_i are different, the propagation of each mass state will be different, and then the presence of any off-diagonal elements in the PMNS

matrix $\mathbf{U}$ can cause the flavour of a neutrino to change during propagation.[3]

In principle the choice of "flavour basis" is somewhat arbitrary, and the matrix could be defined instead in the charged lepton sector; the above notation is however common since charged lepton oscillation is in most cases not experimentally observable, essentially due to the large mass differences for the charged leptons [44]. The PMNS matrix is analogous to the CKM quark mixing matrix, and describes the same underlying physics for the charged and neutral leptons, even though the probing of the matrix elements takes different approaches in the two cases.

This matrix introduces 18 free parameters (9 complex numbers), which the unitarity condition reduces to 9 free parameters. Considering that Dirac fields can be freely redefined by an unobservable complex phase, we could naively further reduce 6 parameters by redefining the charged lepton and neutrino fields – however examining the charged-current Lagrangian $\mathcal{L}_{CC}$, it is seen that one complex phase in the lepton sector must be kept as physically observable [45]. The 3-flavour mixing matrix can thus be fully described by four parameters. A common parameterization uses the Euler mixing angles

$$
\mathbf{U} = \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ \sin\theta_{12} & -\cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix}
$$
$$
\times \begin{pmatrix} \cos\theta_{13} & 0 & e^{-i\delta_{CP}}\sin\theta_{13} \\ 0 & 1 & 0 \\ -e^{-i\delta_{CP}}\sin\theta_{13} & 0 & \cos\theta_{13} \end{pmatrix}, \tag{4}
$$

where the phase δ_{CP} controls the size of the imaginary part of the matrix, and thus controls the level of charge-parity violation (originating in the fact that antiparticle interactions contain the conjugate matrix). In the case of Majorana neutrinos, the neutrino field phases may not be freely redefined due to the condition $\nu = \nu^c$ (where C represents charge conjugation), and two extra phases are needed; these are commonly parameterized as an additional rotation to the above matrix by

$$
\mathbf{U}_{\text{Majorana}} = \mathbf{U} \times \begin{pmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5}
$$

---

3 It is then clear that the lepton family number symmetries present in the SM (L_e, L_μ and L_τ) are not conserved, while the total lepton number L_T is still conserved, assuming a unitary matrix.

1.3.2   *Oscillation*

Oscillation experiments generally measure neutrino interactions in the flavour basis, at some distance L from the initial creation of the neutrino, and we are therefore interested in the probability

$$P(\nu_\alpha \rightarrow \nu_\beta) = \left| \langle \nu_\beta | \nu_\alpha(L) \rangle \right|^2 , \tag{6}$$

where $|\nu_\beta\rangle$ is the state vector for a flavour β and $|\nu_\alpha(L)\rangle$ represents the propagated state vector that was *initially* a flavour α. To calculate $|\nu_\alpha(L)\rangle$, we first consider the propagation of a neutrino mass state ν_i under the assumption of a plane wave solution of definite momentum p and energy E, which can be simply described by the phase factor

$$|\nu_i(L)\rangle = e^{-i(Et - pL)} |\nu_i\rangle . \tag{7}$$

A state of initial flavour α is a weighted superposition of three mass states, as given by the PMNS matrix $\mathbf{U}$, and can thus be propagated as

$$|\nu_\alpha(L)\rangle = \sum_{i=1}^{3} U_{\alpha i} e^{-i(E_i t - p_i L)} |\nu_i\rangle . \tag{8}$$

As a simplifying assumption we note that neutrinos are generally relativistic, so we may approximate $t \approx L$ and furthermore assume that the mass states have similar kinetic (and therefore total) energies $E_i \approx E$. In this case we can expand $p_i = \sqrt{E^2 - m_i^2} \approx E(1 - m_i^2/2E^2)$, and we obtain

$$|\nu_\alpha(L)\rangle \approx \sum_{i=1}^{3} U_{\alpha i} e^{-i(m_i^2 L / 2E)} |\nu_i\rangle . \tag{9}$$

Finally, explicitly expressing the above in terms of the flavour basis β gives

$$|\nu_\alpha(L)\rangle \approx \sum_{\beta=e,\mu,\tau} |\nu_\beta\rangle \langle \nu_\beta| \left[\sum_{i=1}^{3} U_{\alpha i} e^{-i(m_i^2 L / 2E)} |\nu_i\rangle \right] \tag{10a}$$

$$= \sum_{\beta=e,\mu,\tau} |\nu_\beta\rangle \left[\sum_{i=1}^{3} U_{\alpha i} U_{\beta i}^* e^{-i(m_i^2 L / 2E)} \right] , \tag{10b}$$

and any oscillation probability Eq. (6) for a chosen flavour of β may be read off from the bracketed coefficients of Eq. (10b) in terms of m_i, E, L, and U.

In principle, a plane-wave approximation, coupled with an equal-energy (or equal-momentum) approximation for the mass states, may

not be a particularly accurate model. A wave-packet treatment of the oscillation was first suggested as necessary by [46] and performed in [47] by Giunti, Kim, and Lee, who later showed that the single-flavour state vectors are not even themselves well-defined, except in the relativistic limit [48]. Further development of the theoretical framework of oscillation is in fact still under active discussion; a good overview is given in [49]. However, the oscillation probabilities calculated with the simple plane-wave model are consistent with the more comprehensive calculations in the domain $E \gg m$, where neutrino interactions are actually observable, so we may proceed using the formulae as shown above.

Oscillation probabilities are often given depending on the parameterization of the mixing matrix U, and in the three-flavour case can result in quite long formulae. In the two-flavour case for Dirac neutrinos however, the parameterization of U is reduced to only one free parameter θ:

$$\mathbf{U}_{\text{two-flavour}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{11}$$

and the probability can be simply calculated as

$$P(\nu_\alpha \rightarrow \nu_\beta) = \sin^2(2\theta)\sin^2\left(1.27\frac{\Delta m^2\,[\text{eV}^2]\,L\,[\text{km}]}{E\,[\text{GeV}]}\right) \tag{12}$$

where $\Delta m^2 \equiv m_\beta^2 - m_\alpha^2$ and $\alpha \neq \beta$. This simplified yet illustrative case shows the principle of neutrino oscillation measurements; for a known neutrino beam of flavour α and path-length to energy ratio L/E, we may study the neutrino oscillation parameter(s) θ and size of the mass-difference(s) $|\Delta m^2|$, if not the ordering or the absolute value of the neutrino masses. It also illustrates why the atmospheric sample's wide range of path-lengths and neutrino energies make it such a useful neutrino source.

1.3.3 *Matter effects*

The previous section discussed neutrino oscillations in vacuum; the situation will be different in the presence of normal matter, where almost all charged leptons are electrons. While the NC interactions are identical between neutrino flavours, electron neutrinos may undergo CC coherent forward scattering reactions with electrons, of the leptonic form in Fig. 3b where we set $\alpha = \beta = e$; muon neutrinos cannot undergo such a process, and thus the oscillation scheme will be modified (of course, muon neutrinos with $E_\nu > m_\mu$ may still undergo the process where $\alpha = \mu$ and $\beta = e$, but this is not a coherent process). This modification will be shown to be equivalent to a change in the *effective masses* of the propagating neutrinos, and is generally

called the *matter effect* [50]; a somewhat appropriate analogy would be to consider the dispersion of a light beam inside a material with a refractive index.

The coherent scattering interaction for Dirac neutrinos introduces a term of the form

$$\mathcal{H}_{\text{mat}} = \frac{G_F}{\sqrt{2}}\, \bar{\nu}_e \gamma^\alpha (\mathbf{1} - \gamma^5)\nu_e\, \bar{e}\gamma_\alpha(\mathbf{1} - \gamma^5)e, \tag{13}$$

which can be simplified by applying the chiral projection operator, notating left-handed chirality by the subscript *L*, to

$$\mathcal{H}_{\text{mat}} = \sqrt{2}G_F\, \bar{\nu}_{e,L}\gamma^\alpha \nu_{e,L}\, \bar{e}_L\gamma_\alpha e_L. \tag{14}$$

Further considering an interaction in the rest frame of unpolarized normal matter, only the γ^0 contribute for the electrons, which gives

$$\mathcal{H}_{\text{mat}} = \sqrt{2}G_F\, \bar{\nu}_{e,L}\, \nu_{e,L}\bar{e}_L e_L \tag{15}$$

which is simply a potential energy term for electron neutrinos, with the potential

$$\mathcal{V}_{\text{mat}} = \sqrt{2}G_F N_e \tag{16}$$

where N_e is the electron density.

Looking back to Sec. 1.3.2, putting our description of neutrino oscillation into matrix form, and differentiating with respect to time (for $t = L$) we have

$$i\partial_t \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} = \frac{1}{2E} \begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3 \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{17}$$

which has the same form as the time-dependent Schrödinger equation $i\partial_t |\nu_i\rangle = \mathcal{H}_{\text{vac}} |\nu_i\rangle$, where $\mathcal{H}_{\text{vac}}$ can be interpreted as the vacuum Hamiltonian. However, we must be careful to not over-interpret the similarity, as we are explicitly dealing with *relativistic* neutrinos. Nonetheless, under the same approximations used in Sec. 6, it is seen that [51, 52] the effective Hamiltonian for propagating vacuum and matter effects is simply $\mathcal{H}_E = \mathcal{H}_{\text{vac}} + \mathcal{H}_{\text{mat}}$ and the propagation equation is then

$$i\partial_t \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \left[\mathbf{U}\frac{1}{2E}\begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3^2 \end{pmatrix}\mathbf{U}^\dagger + \begin{pmatrix} \mathcal{V}_{\text{mat}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}. \tag{18}$$

Here, the mass matrix was moved to the flavour basis, while the matter effects are of course native to the flavour basis; it is then clear that

the MSW effect can be seen as a change of the effective masses of the neutrinos. In general this matrix can be simplified to absorb an unobservable change to the overall phase of all three flavours, for example a subtraction involving m_1^2, leaving only Δm_{12}^2 and Δm_{13}^2, and in fact the NC potential has already implicitly been absorbed. After re-diagonalization of the matrix, the oscillation formulae such as Eq. (12), can be given in matter in terms of two effective mass differences $\Delta m_{ij,\text{eff}}^2$ and three effective mixing angles $\theta_{ij,\text{eff}}$ (plus the complex phase).

For a concrete example, we assume ν_τ does not exist, and take the two-flavour case of ν_e and ν_μ with notation as in Eqs. (11) and (12). Here, the effective Hamiltonian becomes

$$\mathcal{H}_E = \frac{1}{4E} \begin{pmatrix} -\Delta m^2 \cos(2\theta) + 2E\mathcal{V}_{\text{mat}} & \Delta m^2 \sin(2\theta) \\ \Delta m^2 \sin(2\theta) & \Delta m^2 \cos(2\theta) - 2E\mathcal{V}_{\text{mat}} \end{pmatrix}, \quad (19)$$

where a constant term was removed and the $\mathcal{V}_{mat}$ split across the diagonal by a redefinition of the overall phase. We can always diagonalize this matrix to

$$\mathcal{H}_M = \frac{1}{4E} \begin{pmatrix} -\Delta m_{\text{eff}}^2 & 0 \\ 0 & \Delta m_{\text{eff}}^2 \end{pmatrix} \quad (20)$$

by finding a transformation $\mathcal{H}_M = \mathbf{U}_M \mathcal{H}_E \mathbf{U}_M^\dagger$, where

$$\mathbf{U}_M = \begin{pmatrix} \cos\theta_{\text{eff}} & \sin\theta_{\text{eff}} \\ -\sin\theta_{\text{eff}} & \cos\theta_{\text{eff}} \end{pmatrix}, \quad (21)$$

and thus we have the new effective parameters. Compared to the vacuum parameters, the new mixing angle will be given by

$$\tan(2\theta_{\text{eff}}) = \tan(2\theta) / \left(1 - \frac{2E\mathcal{V}_{\text{mat}}}{\Delta m^2 \cos(2\theta)} \right). \quad (22)$$

Considering the various cases of whether or not $\cos(2\theta) > 0$ and (or) $m_2 > m_1$, the effective mixing strength can be increased or decreased given an appropriate material $\mathcal{V}_{\text{mat}}$. In particular, if both are positive we have the possibility of a resonant enhancement, where the bracketed term becomes zero and the effective mixing angle becomes maximal ($\theta_{\text{eff}} = 45°$); considering Eq. (12), depending on the path-length and energy, complete conversion between flavours would be possible for any intrinsic value of θ.

Above, we assumed that N_e was constant. In cases where the observed neutrinos pass through areas of varying electron density, such as in the Sun or Earth, generally division of the density profile into various zones or a numerical treatment can be used. A more careful analytical treatment of propagation through the Sun, during which the resonant enhancement point is gradually crossed, showed that total conversion of the ν_e flux (created in the fusion process) to ν_μ is
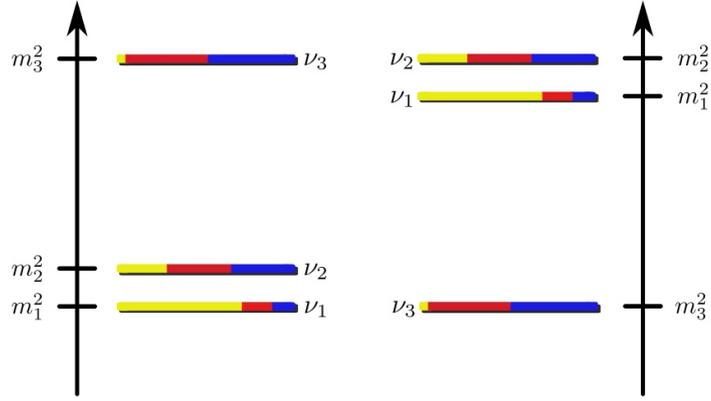
Figure 4: The "normal" (left) and "inverted" (right) possible hierarchies for neutrinos. The three colors show the flavour components of each mass state (e, μ, τ). [55]

possible for all neutrinos above a certain energy; this enhancement is known as the *MSW effect*, after the first published authors Wolfenstein (on matter effects) [50] and Mikheyev and Smirnov (on the resonant enhancement) [53]. A particularly clear description of the effect, including a correction of the sign of the potential, is given by Bethe in [54].

Finally, it should be noted that importantly for anti-neutrinos, the potential $\mathcal{V}_{\mathrm{mat}}$ will have the opposite sign and thus quite different matter effects. The extra phases of Majorana neutrinos however can not change the effective parameters.

To summarize, unlike the vacuum oscillation probabilities, probing neutrino oscillations in matter with neutrinos and anti-neutrinos can in principle show different observable effects depending on the neutrino mass hierarchies and mixing parameter octants.

1.3.4   *Current experimental status*

Currently all the mixing angles θ_{12}, θ_{23}, and θ_{13} are experimentally measured, and are considerably larger than the mixings seen in the quark sector; however there is still ambiguity as to the true octant of θ_{23} (greater or less than $45°$). The charge-parity violation phase has not yet been significantly measured as non-zero. Both mass splittings Δm_{12}^2 and Δm_{23}^2 are measured, with $\Delta m_{23}^2 \gg \Delta m_{12}^2$. The mass hierarchy (i.e. ordering of the masses) is known at least for Δm_{12}^2 where $m_2 > m_1$, however the sign of Δm_{23}^2 is not yet known, as illustrated in Fig. 4.

Constraints on the absolute neutrino mass scale have been obtained by methods including direct travel-time measurements from the Supernova 1987 [40], cosmological energy-density constraints and double-beta decay experiments [56]. The discovery of Majorana neutrinos (the existence of which is widely believed in the theoretical commu-

PARAMETER	NOMINAL VALUE	UNCERTAINTY
Δm_{21}^2	7.50×10^{-5}	$\pm 0.2 \times 10^{-5}$
$\lvert \Delta m_{32}^2 \rvert$	2.32×10^{-3}	$\pm 0.12 \times 10^{-3}$
$\sin^2 \theta_{12}$	0.310	$-0.014 + 0.023$
$\sin^2 \theta_{23}$	0.500	± 0.112
$\sin^2 \theta_{13}$	0.0240	± 0.0023
δ_{CP}	0	$0 \to 2\pi$
sign of Δm_{32}	$+$	$+$ or $-$

Table 1: Nominal value and $1\,\sigma$ uncertainty (or for the CP violation phase and mass hierarchy, possible values) of the oscillation parameters used in this thesis, based on [59]. The unit of Δm_{ij}^2 is eV^2.

nity) may be possible by detection of neutrinoless double-beta decay [57]. The existence of Dirac neutrinos is currently believed to be not directly provable, as it would require an infinitely stringent limit on such a decay.

It is predicted that the sign of Δm_{23}^2 and the θ_{23} octant may be probed in the next generation of neutrino detectors [17] via atmospheric neutrinos passing through the Earth and consideration of their matter effects, as such detectors are expected to have a stronger separation of neutrinos and anti-neutrinos. The δ_{CP} phase is similarly expected to be measurable, for a significant fraction of the possible range, by long baseline neutrino beam experiments. The existence of matter effects, although an integral part of the standard oscillation theory and confirmed with high significance (if somewhat indirectly) in the solar neutrino model, should also be observable inside of the Earth [58] as a further check on the consistency of the theory.

A summary of current measurements is given in Tab. 1.

1.4 MASSIVE NEUTRINOS

Since the discovery of neutrino oscillations, the understanding of the origin of neutrino mass has become an important task in theoretical and experimental physics. Here the principles are discussed briefly, as while essential for a general understanding of neutrino physics, the results here are not directly used in later parts of this thesis.

1.4.1 *Mass terms*

In order to incorporate neutrino masses into the SM, we will see that some new fields must always be postulated. The standard Dirac mass terms for neutrinos (assuming one generation) take the form

$$-m_D \bar{\nu}\nu = -m_D(\bar{\nu}_L \nu_R + \bar{\nu}_R \nu_L). \tag{23}$$

FIELD	I	I_3	Y	Q
$\begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$	$\frac{1}{2}$	$\frac{1}{2}$ $-\frac{1}{2}$	-1	0 1
e_R	0	0	-2	-1
ν_R	0	0	0	0
$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\frac{1}{2}$	$\frac{1}{2}$ $-\frac{1}{2}$	$\frac{1}{3}$	$\frac{2}{3}$ $-\frac{1}{3}$
u_R	0	0	$\frac{4}{3}$	$\frac{2}{3}$
d_R	0	0	$-\frac{2}{3}$	$-\frac{1}{3}$
$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$	$\frac{1}{2}$	$\frac{1}{2}$ $-\frac{1}{2}$	1	1 0

Table 2: Charges of the SM fields (first generation), and the necessary charges of a sterile neutrino singlet ν_R. Here ϕ is the Higgs doublet, and the convention is $Q = I_3 + \frac{Y}{2}$, where Q is electric charge, I is isospin, and Y is weak hypercharge. The leptons also carry lepton number $L_T = 1$, except in the case of Majorana neutrinos where $L_T = 0$.

The immediate problem is that these mass terms are chiral, whereas in the SM no right-handed neutrinos (or left-handed anti-neutrinos) exist by default, since they were not observed to participate in the weak interaction. It is of course possible to add a right-handed neutrino field, as long as it belongs to an isospin singlet, i.e. a field that does not experience any forces except gravity – often referred to by such terms as "sterile neutrino" or "neutral heavy lepton". The terms on the right side of Eq. (23) then have $I_3 = \pm\frac{1}{2}$ and $Y = \pm 1$ (see Tab. 2), and can be written into the SM by coupling with the standard Higgs' doublet Φ such that the SM charges are conserved. The total lepton number is also explicitly conserved by the $\bar{\nu}\nu$ pairing.

Another possible mass term, though not used in the SM, are the Majorana mass terms with the form

$$-\frac{1}{2}m_M^L(\bar{\nu}_L\nu_L^C + \bar{\nu}_L^C\nu_L) \tag{24a}$$

$$-\frac{1}{2}m_M^R(\bar{\nu}_R\nu_R^C + \bar{\nu}_R^C\nu_R) \tag{24b}$$

where the charge conjugation operator C is essentially similar to overbar, but maintains the necessary row or column structure. Majorana mass terms were originally designed to be used in the case of Majorana fermions – particles that are their own anti-particles, i.e. $\nu = \nu^C$. Naively this kind of term could avoid the problem of extra neutrino fields, if we take only the terms in Eq. (24a) as physical; however,

these terms have $I_3 = \pm 1$ and $Y = \pm 2$, and thus would require an extension[4] of the Higgs sector to include a field with $Y = 2$. Lepton number conservation is also explicitly broken for a Dirac fermion, and even for a Majorana fermion we see that there is no U(1) symmetry $\nu_L \rightarrow e^{i\varphi}\nu_L$, and by Noether's theorem no associated conserved current.[5]

1.4.2 *The seesaw mechanism*

It is also possible that *both* types of mass terms are actually present, and write all of the above terms into a matrix

$$\mathcal{L}_{\mathrm{m}} = -\frac{1}{2} \begin{pmatrix} \bar{\nu}_L & \bar{\nu}_R^C \end{pmatrix} \begin{pmatrix} m_M^L & m_D \\ m_D & m_M^R \end{pmatrix} \begin{pmatrix} \nu_L^C \\ \nu_R \end{pmatrix} + \text{h.c.,} \qquad (25)$$

in which case it seems that the right-handed fields (ν_R and ν_L^C) and left-handed fields (ν_L and ν_R^C) no longer have definite masses individually. The matrix can be diagonalized to

$$\mathcal{L}_M = -\frac{1}{2} \begin{pmatrix} \bar{\nu}_a & \bar{\nu}_b \end{pmatrix} \begin{pmatrix} m_a & 0 \\ 0 & m_b \end{pmatrix} \begin{pmatrix} \nu_a \\ \nu_b \end{pmatrix} + \text{h.c.,} \qquad (26)$$

where the a and b neutrinos must have the structure e.g. $\nu_a = \nu_{a,L} + \nu_{a,L}^C$; in other words, if both mass terms are present the neutrinos must be Majorana fermions.

The observable masses m_a and m_b then depend on the physically-motivated choices we make for m_M^L, m_M^R, and m_D. As discussed in Sec. 1.4.1, m_M^L is forbidden in the SM. The Dirac mass m_D is forbidden before symmetry breaking but allowed after symmetry breaking, thus should take its value from the Higgs mechanism with the vacuum expectation value of 246 GeV. Finally, m_M^R is generally allowed in the SM, and could be generated by any mechanism, including new physics at the Grand Unified Theory (GUT) scale, which is plausibly around 10^{15} GeV.

Then, assuming the Yukawa couplings are of the order unity, we can take

$$m_L^M = 0, \quad m_D \sim 10^2 \text{ GeV}, \quad m_L^R \sim 10^{15} \text{ GeV}, \qquad (27)$$

which by examination of the matrix diagonalization gives

$$m_a \sim 10^{-4} \text{ eV}, \quad m_b \sim 10^{15} \text{ GeV}. \qquad (28)$$

This choice of masses is known as the *seesaw mechanism* [60], which may provide some justification for the comparatively tiny observed

---

4 This could also be expressed by a non-renormalizable dimension 5 operator within the SM, although this explicitly turns the SM into an effective field theory.

5 However, the violations of L_T in the overall Lagrangian will be small; this does however allow the previously-mentioned double-β decay.

neutrino masses $m_i \sim m_a$. The name is also applied to various other models and choices of the parameters that can result in the observed hierarchy.

In the above case, as can be seen from the mixing of the doublet ν_L and singlet ν_R into ν_a and ν_b, the sterile state participates in the neutrino flavour mixing.[6] The matrix **U** must then be extended to include the sterile flavour(s); at the present time, the measurements of the matrix parameters in the three-flavour scheme are consistent with the unitarity of the matrix, and there is no conclusive evidence for the existence of steriles (although there is some tension between experimental results [61]). In any case, the oscillation effects of sterile neutrinos must be small, and while the discussion of neutrino mass is of great interest, the effects will be ignored for the rest of this thesis.[7]

---

6 In the three-flavour case this is generally true for Majorana neutrinos, but not Dirac ones.

7 For a full discussion see e.g. [51].

THE ATMOSPHERIC NEUTRINO FLUX

2.1 ORIGIN

The atmospheric neutrino flux originates from the cosmic ray flux colliding with the Earth's atmosphere. The cosmic ray flux consists mostly of protons and Helium nuclei in the ratio 9 : 1, with heavier nuclei at the $\sim$1% level, and is isotropic within 0.1% [62]. Electrons, positrons and electromagnetic radiation make up an even smaller percentage, negligible from the neutrino physics perspective. The cosmic rays are believed to originate mostly from outside the Solar System, from sources such as supernova remnants in our galaxy and the active nuclei of other galaxies; the proportional contribution of the sources, and exact nuclei composition across the wide energy range, is however still under study (e.g. by [63]).

In general cosmic ray collisions in the atmosphere create particle showers which are dominated by light mesons such as pions and kaons. A graphical representation of such a shower is shown in Fig. 5. The dominant decay chain (with branching ratio > 99%) of charged pions involves the creation of neutrinos as

$$\pi^{\pm} \to \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \tag{29}$$

and the muons further decay via the "Michel decay" as

$$\mu^{\pm} \to e^{\pm} + \nu_e(\bar{\nu}_e) + \bar{\nu}_{\mu}(\nu_{\mu}). \tag{30}$$

Charged kaons undergo the identical decay as Eq. (29) with a branching ratio of 64%, with the remainder of decays being either hadronic (to pions) or semileptonic as

$$K^{\pm} \quad \to \quad \pi^0 + e^{\pm} + \nu_e(\bar{\nu}_e) \tag{31a}$$

$$K^{\pm} \quad \to \quad \pi^0 + \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \tag{31b}$$

with branching ratio 5% and 3% respectively. The neutral pion decays generally as $\pi_0 \to \gamma + \gamma$, and the neutral kaon (via the short and long weak eigenstates) decays primarily to pions. Therefore, considering that the majority of decays to neutrinos will be by charged pions or kaons as in Eq. (29), then by muons as in Eq. (30), we can make the first-order approximations for the neutrino flux

$$\frac{\nu_{\mu} + \bar{\nu}_{\mu}}{\nu_e + \bar{\nu}_e} \quad \sim \quad 2 \tag{32a}$$

$$\frac{\nu_{\mu}}{\bar{\nu}_{\mu}} \quad \sim \quad 1 \tag{32b}$$

$$\frac{\nu_e}{\bar{\nu}_e} \quad \sim \quad \frac{\mu^+}{\mu^-} \tag{32c}$$
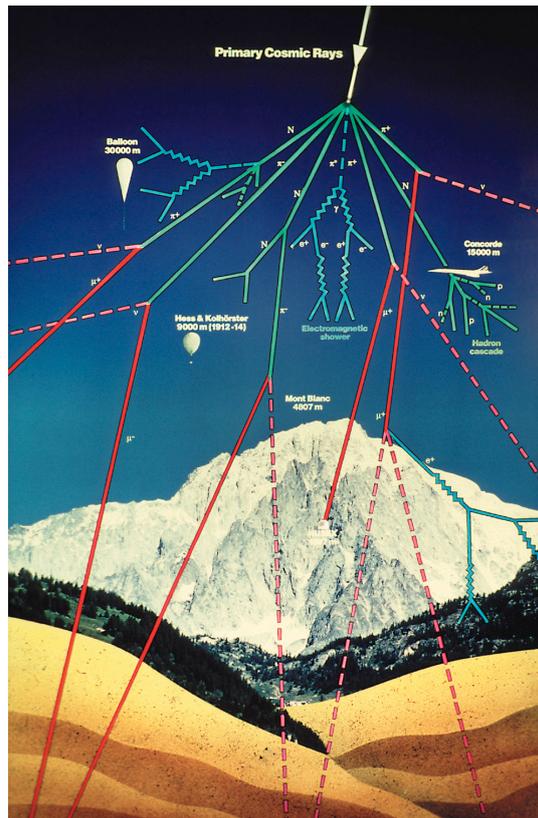
Figure 5: A schematic representation of a primary cosmic ray collision and shower in the atmosphere. [65]

assuming that all decay chains reach their final states. More accurate determinations of the flux, such as the overall normalization, must be based on a more complete study of the primary cosmic ray flux, the hadronic processes in the atmosphere, and the geomagnetic field effects.

At energies above 100 TeV, a different mechanism than the above is expected to dominate, whereby neutrinos are primarily created from the decay of heavier mesons involving charm quarks; these are termed "prompt" neutrinos [64]. While the prompt flux is expected to be low, it is in principle distinguishable by large-volume neutrino telescopes [41]. However, the Super-Kamiokande detector does not have significant power to observe such a flux, and this thesis focuses on the "conventional" neutrino flux from pions and kaons.

2.2   ATMOSPHERIC NEUTRINO FLUX SIMULATIONS

2.2.1   *Historical approach*

As mentioned in the above chapter, any simulation of the atmospheric neutrino flux must generally take into account the primary cosmic ray spectra, all hadronic effects and interactions in the atmosphere

(in other words, the yield of secondary particles), and the effects of the geomagnetic field. At the simplest level the neutrino flux as a function of energy $\phi_\nu(E)$, for an infinitesimal solid angle, could be produced by a simple convolution

$$\phi_\nu(E) = \sum_N \int_0^\infty \phi_N(E' - E) \times Y_N(E', E) \times R_N(E') \, dE' \qquad (33)$$

where N is a sum over nuclei (from protons to helium and heavier nuclei), $\phi_N(E')$ is the primary cosmic ray spectrum, $Y_N(E', E)$ represents the average yield of neutrinos reaching the ground for a nucleon N of energy E, and finally $R_N(E')$ is a simple energy-filtering Heaviside step function (i.e. zero or one). The energy cutoff of the step function is calculated based on the incident location and direction of the primary cosmic ray, and decides whether a nucleon could have reached the atmosphere or not depending on the magnetic field effect known as the *rigidity cutoff* (explained in Sec. 2.2.2).

In this method, the distributions are calculated choosing a particular point on the Earth's surface, and integrating around a 4π solid angle. While this may be possible analytically, the disadvantage of such a method is obvious; while the geomagnetic cutoff information is included for the primary particle, all secondary particles are assumed to travel in a perfectly straight line, ignoring geomagnetic effects, and even conservation of momentum of the decay products. Furthermore, only the average neutrino yield is used. In order to recreate the neutrino flux accurately, a fully three-dimensional Monte Carlo simulation, with individual tracking of all particles, seems necessary.

2.2.2 *Rigidity cutoff*

To understand the meaning of the rigidity cutoff, consider a random incident location and direction of a cosmic ray collision on the atmosphere. After solving the equations of motion in the geomagnetic field in the backwards time direction, we may find that the tracks below a certain energy range actually originated on the Earth's surface; such tracks are forbidden as they imply that the cosmic ray would have originally had to travel through the Earth. An example tracing of two tracks for a fixed energy is shown in Fig. 6, where only one track passes the rigidity cutoff. For a given ground location, we can draw a map of the rigidity cutoff as a function of direction and primary cosmic ray energy, such as the one shown in Fig. 7; the irregular shape originates in the fact that the geomagnetic field is not perfectly symmetric, but contains some complex structure.
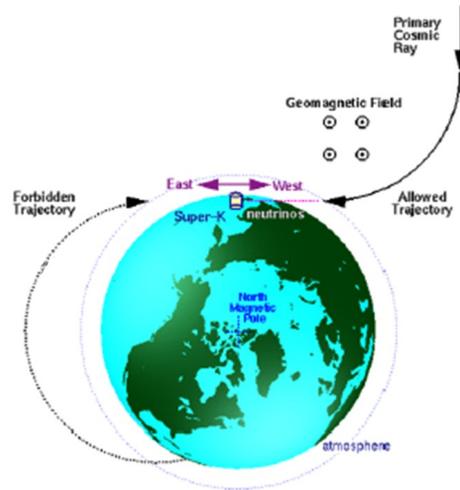
Figure 6: Illustration of the rigidity cutoff effect, for two primary cosmic ray collisions that could result in neutrino events at Super-Kamiokande. After tracing backwards in time, only one cosmic ray passes the rigidity cutoff test. [66]
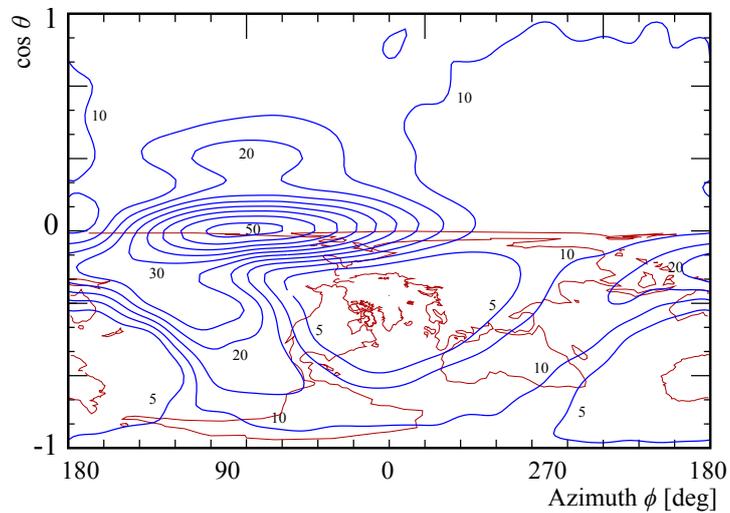
Figure 7: Contour map of the rigidity cutoff, seen from the perspective of the SK detector. The labeling of the azimuth angle ϕ and zenith angle θ is the same as in Sec. 6.2. The units represent the rigidity cutoff in GV (e.g. for a proton with a charge of $+e$, this is the required energy in GeV to pass the cutoff). [67]
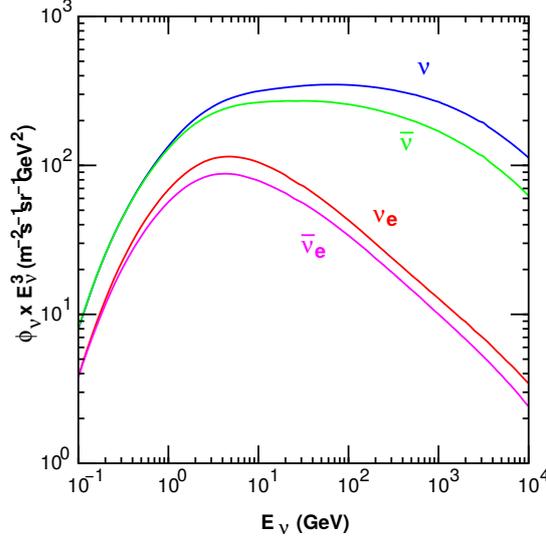
Figure 8: The energy spectra of the neutrino fluxes at the Super-Kamiokande site as predicted by the HKKM15 model. [68]

2.3 THE HKKM FLUX MODEL

2.3.1 *Current model*

The default flux model used at Super-Kamiokande is produced by the HKKM group [18, 19, 68], and has been in continuous development since 1988. The model tracks the neutrino flux at specific locations on the Earth corresponding to the current neutrino detectors, such as Super-Kamiokande, INO, IceCube, and so on. For each location, the energy and directional spectra of the neutrino flux are published in table format, along with instructions for accurate interpolation. An example of the final energy spectrum of the fluxes is shown in Fig. 8. In this section, the current status and approach of the most recent flux model (2015, [68]) will be discussed.

2.3.2 *Three-dimensional calculation*

The calculation is fully three-dimensional, in general tracking every necessary particle from its generation to interaction within the atmosphere. The Earth is assumed as a sphere of radius $R_e = 6378.140$ km (while this is not exactly true, the approximation is expected to have a negligible effect on the final neutrino flux), and furthermore the *injection* sphere is defined at $R_e + 100$ km, and the *escape* sphere at $10R_e$. For each cosmic ray simulation, the energy and nuclei size is sampled from the data-based primary cosmic ray distributions taken from the AMS [24, 25] and BESS [22, 23] groups, and a random position and direction on the injection sphere is chosen. A rigidity cutoff test, as explained in Sec. 2.2.2, is performed by tracing the equations

of motion in the IGRF [69] magnetic field model backwards in time, until the track hits either the escape sphere or returns to the injection sphere. If the particle passes this test of extraterrestrial origin, it is fed into an atmospheric propagation simulator.

The propagation simulator follows the traditional cascade model, tracking the motion of all particles step-by-step, accounting for the forces applied by the magnetic field. For example, the bending of each muon's path before decay has a significant effect on the flux at a given location, compared to a simple one-dimensional model.

All appropriate particle decays are implemented. The hadronic interactions with air nuclei are handled by a combination of the JAM model (a subset of the PHITS Particle and Heavy-Ion Transport System code [70]) at $E < 32$ GeV, and a modified DPMJET-III [71] code above that, which was updated to give better agreement to the atmospheric muon spectra of secondary cosmic rays measured by the BESS group. The probabilities for interactions are primarily based on the air density, and to a lesser degree temperature, both taken from the NRLMSISE-00 [72] model (recently updated from the US-standard '76 model [73]). The model is based on empirical measurements and contains a full positional (latitude, longitude, and altitude) and seasonal dependence from the Earth's surface far into space. The seasonal dependence is largest at the poles with a maximum $\sim 10\%$ variance in the ν_e and $\bar{\nu}_e$ fluxes (less for ν_μ and $\bar{\nu}_\mu$) at around $E_\nu = 3.2$ GeV, although quickly decreases moving towards the equator where the variance is $< 1\%$.

Finally, the simulation records the neutrino flux arriving at particular areas of the Earth's surface, including of course the flux that passes through from opposite sides of the Earth, corresponding to the position of the various neutrino detectors. However, the actual size of the detectors is tiny compared to the size of the Earth, and thus when using the fully three-dimensional calculation, even with extremely high statistics simulations it is unlikely that many neutrinos will directly impact the detectors. The concept of "virtual detectors" is then introduced, which are areas covering the Earth's surface up to a radius of 1113.2 km from the detector positions (which corresponds to a change in longitude of 10° at the equator), and any neutrino passing through these areas is registered. This of course requires a correction factor depending on the zenith angle, with a higher weighting for events arriving horizontally at the detector. While the virtual detector can introduce some bias due to essentially averaging the fluxes over a wide area, the strength of this bias can be examined by testing virtual detectors of various size at a given location. After application of various correction factors, the bias is reduced to the 1% level.
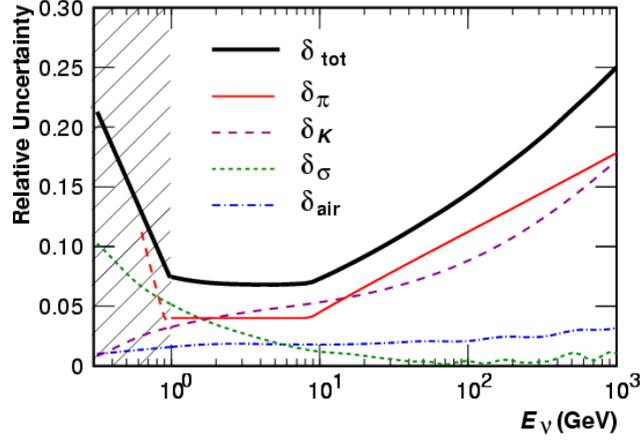
Figure 9: The estimated systematic errors on the HKKM07 model. The error sources are as follows: atmospheric density profile uncertainties δ_{air}, mean free path (cross-section) uncertainties δ_σ, kaon flux uncertainties δ_K, pion flux uncertainties δ_π, and total uncertainty δ_{tot}. [75]

2.3.3 *Systematic error estimations*

An estimation of the systematic uncertainties on the HKKM07 simulation, produced by the HKKM group, is shown in Fig. 9. It can be seen that the flux is estimated very accurately in the $1 \sim 10$ GeV range, but with increasing uncertainties at energies down towards the MeV scale or up towards the TeV scale. The π and K flux uncertainties mainly come from experimental uncertainties on various μ flux measurements, from which the π/K ratio and normalizations are extracted. In the HKKM11 model, the uncertainty is expected to be reduced by $\sim 5\%$ at < 1 GeV due to the updates to the hadronic interaction simulation, and for the HKKM15 model some reduction in the atmospheric density profile uncertainty is also expected [74].

Part II

THE EXPERIMENT

I think I'm thinking, therefore I possibly might be.
— Arnold J. Rimmer

THE SUPER-KAMIOKANDE DETECTOR

3.1 PRINCIPLES OF CHERENKOV DETECTORS

3.1.1 *Cherenkov radiation*

The basic principle of Cherenkov detectors is to observe the Cherenkov radiation produced when a charged particle moving faster than the phase velocity of light, in a particular medium, emits electromagnetic radiation around the visible range. This emission is due to the polarization and relaxation of the material close to the particle track, generating wavelets at each point, which cause constructive interference in a forward direction as visualized in Fig. 10.

Purified water can provide at low cost a massive target volume, while having a long light attenuation length, allowing its use in large-volume Cherenkov detectors. The speed of light in a medium is given by $v_\gamma = c/n$, where the refractive index of water n is approximately 1.33. Therefore, considering the relative mass of a given type of particle, the minimum energy required for them to exceed v_γ and thus become observable by Cherenkov light may be easily calculated, with some commonly observed particles shown in Tab. 3 (although in practice, the observational threshold of a given detector will be somewhat higher than these values, considering light collection efficiency and background noise).

The Cherenkov light emission occurs at the Cherenkov angle ϑ, depending on the velocity of the charged particle v_p as $\cos\vartheta = c/(nv_p)$. For particles with an energy precisely at the Cherenkov threshold, we have $\vartheta = 0°$, and then ϑ increases with higher particle energies; for relativistic particles with $v_p/c \approx 1$ in water, we have $\vartheta = 42°$. The Cherenkov angle is thus useful to differentiate low-energy particles, but loses its effectiveness once particles have total energy just a few times larger than their rest mass, where v_p/c approaches unity.
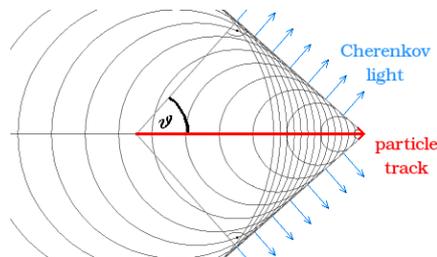
Figure 10: Cherenkov light production by a charged particle track. Dotted angles represent right angles. [55]

PARTICLE TYPE	CHERENKOV THRESHOLD [MEV]
$e^{\pm}$	0.77
$\mu^{\pm}$	159.8
$\pi^{\pm}$	210.7
$p^{\pm}$	1419.0

Table 3: The Cherenkov thresholds (minimum total energy required to emit Cherenkov light) for various particle types in water with $n = 1.33$.

Considering a particle traveling on a short straight track, Cherenkov light will be emitted in a hollow forward-facing cone, that when projected onto a flat surface (e.g. at the edge of the target volume) produces a ring shape. The longer the particle track, the thicker the ring will be, where a track that reaches to the projection surface producing a completely filled ring. For minimally ionizing particles, such as muons with energy $O(\text{GeV})$, the track length and thus particle momentum will be well correlated with the thickness of the ring. For particles such as electrons, which can undergo multiple scattering and electromagnetic showering, the ring shape will be more diffuse and the energy is better correlated with a simple count of the emitted photoelectrons [4].

3.1.2 *Photomultiplier tubes*

Considering the physics of Cherenkov light emission as discussed in the previous paragraph, it is clear that the projection of the Cherenkov light at the boundary of the interaction medium in principle contains information that allows us to reconstruct the direction, energy, and particle type of a charged particle track. A common method of sensitive detection of light is by the use of photomultiplier tubes (PMTs), the basic operating principle of which are shown in Fig. 11. They use the energy of an incident photon to liberate an electron via the photoelectric effect, where the material of the photocathode is chosen as to be photosensitive to the light wavelengths of interest. The probability of emission depending on the absorbed wavelength is known as the quantum efficiency (QE), and an example of a QE distribution for a PMT optimized for detecting Cherenkov light is given in Fig. 12. The electron is then accelerated in an electric field, and amplified via collision with a series of dynodes, eventually producing a current large enough to feed into standard electronics systems. If multiple photons impact the PMT over a short time frame, the output will be a higher total charge, which after calibration can be converted to a count of photoelectrons n_{pe}. The conversion usually cannot be perfect, due to a non-zero width of the output charge distributions for a true impact
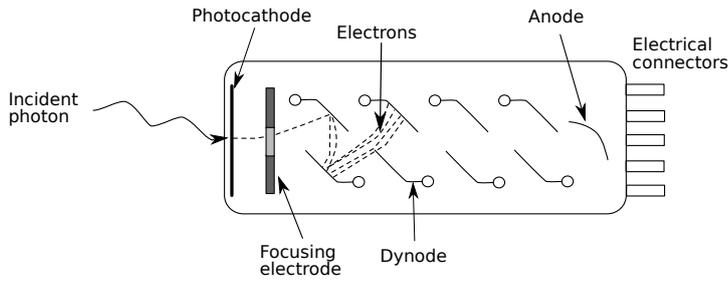
Figure 11: Basic operating principle of the photomultiplier tube. [55]
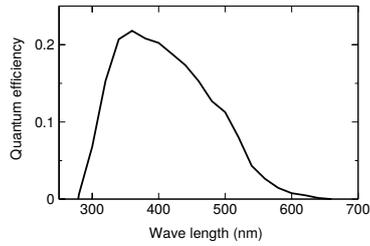
Figure 12: Example distribution of the quantum efficiency of a PMT photo-
cathode (used in Super-Kamiokande). [4]

of a particular number of n_{pe}; an example of PMT charge outputs is
given in Fig. 13 for a high quality modern PMT.
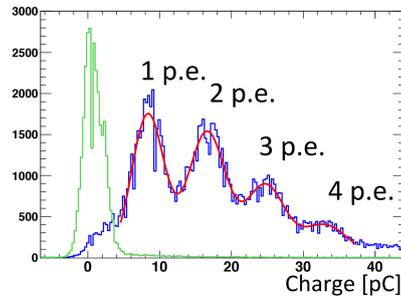
Figure 13: Example raw charge outputs from a modern PMT, showing "dark
rate" data (green), charge outputs during exposure to a light
source (blue), and fitted n_{pe} distributions (red). The dark rate
refers to random outputs in the absence of a light source, coming
from e.g. thermionic emission. Shown for illustrative purposes
only (not used in Super-Kamiokande). [76]
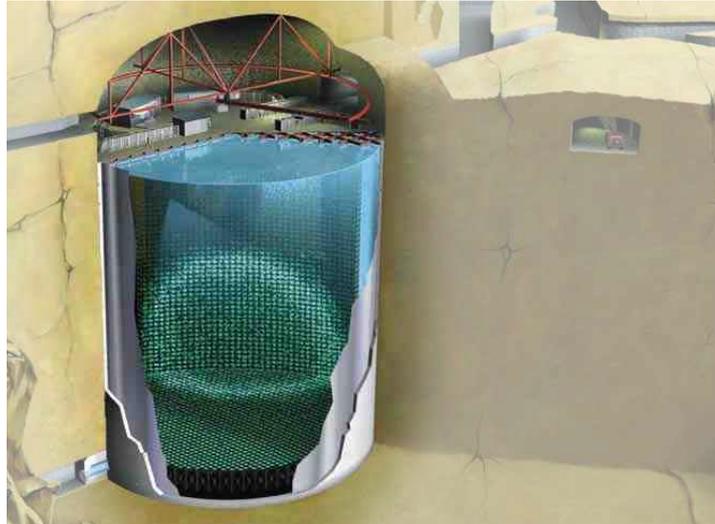
Figure 14: 3D model of Super-Kamiokande. [66]

3.2   STRUCTURE OF THE DETECTOR

Super-Kamiokande (SK) is a water Cherenkov detector located in the Mozumi mine of Gifu prefecture, Japan [77]. A model of the detector is shown in Fig. 14. It is located within Mt. Ikenoyama, with an average rock overburden of ~ 1000 m and density of $2.7\,\mathrm{g\,cm^{-3}}$, in order to minimize the cosmic ray muon flux reaching the detector. About 11,000 20 inch PMTs are mounted on the wall of the detector facing inwards, covering the region known as the inner detector (ID).[1] An optically-separated region on the outer side contains about 2,000 PMTs, known as the outer detector (OD).[2] The main purpose of the OD is to act as a veto for incoming cosmic ray events, however it can also provide further information on charged particle tracks that exit the ID, and muon tracks that enter the detector from below (which may be identified as originating from a neutrino interaction).

Both ID and OD regions are filled with ultra-pure water, approximately 50 kton total, and the region within a boundary 2 m inside the inner wall is considered as the fiducial volume (FV), which contains 22.5 kton of water. This region is defined so that there is an extra buffer against low-energy radioactivities in the surrounding rock and entering cosmic ray particles, which may be falsely reconstructed as contained events, and to exclude tracks passing close by or through the ID PMTs, which are difficult to reconstruct accurately.

The continuous purification of the water is necessary in order to remove radioactive isotopes which may act as a background (especially for solar neutrino measurements, which have energies E_ν within tens of MeV), and to maintain the longest possible light attenuation length.

---

1 The ID uses 20-inch diameter PMTs developed by Hamamatsu Photonics [78].
2 The OD also uses Hamamatsu PMTs, of a different and smaller type, originally taken from the IMB experiment and later partially replaced with a new model [77].
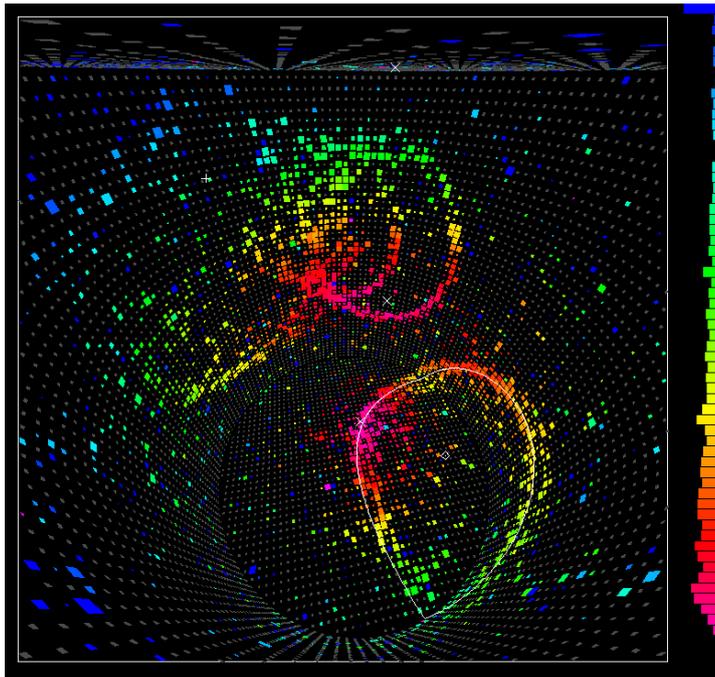
Figure 15: A three dimensional projection of a multi-ring data event. The color scale represents time over an 80 ns period, and charge is represented by the size of the square placed at each PMT position. [66]

A statistic commonly used inside the collaboration is the "light left after 15 m", which for Cherenkov wavelengths is kept at the 90% level. Above the OD at the very top of the tank, there is a cushion of air between the water and the ceiling, which is also filtered in order to reduce the concentration of radon isotopes that seep into the mine air from the rock (despite the rock surfaces being coated with a protective substance). As PMTs depend on an electric potential to accelerate and collect electrons, it is clear that a magnetic field can affect the collection efficiency; in fact, even an 10 mG field would alter the efficiency by approximately 1% [79]. Thus the geomagnetic field must be compensated for to reduce such an effect, which is done by Helmholtz coils placed around the tank.

3.2.1 *Detector phases*

SK has four experimental phases so far. These are designated as SK-I (1996-2001), SK-II (2002-2005), SK-III (2006-2008), and SK-IV (2008-present). The major detector changes that distinguish these periods are as follows.

- **SK-I:** The initial period, ended when an implosion accident destroyed about half of the ID PMTs.

- **SK-II:** The remaining PMTs were rearranged for even coverage (a reduction in total photo-sensitive surface coverage from 40% to 19%), and protective PMT shieldings were introduced.

- **SK-III:** Coverage was restored to the SK-I level with newly produced PMTs.

- **SK-IV:** An improved electronics system was installed, allowing for continuous data recording and thus better reconstruction of sub-events [80].

3.2.2 *Monitoring systems*

The SK detector is ideally in operation 24 hours per day, and several automated and manual processes regularly check the status and condition of the detector. At least one physicist is on duty at all times, in either the mine itself or the nearby campus. Physical properties such as the electronics temperature, water level and temperature, high-voltage power supply levels, and so on, are automatically monitored, as is the data acquisition and processing stream. Any problems are automatically reported to the shift physicist, who also performs additional tasks such as visually monitoring event triggers in real time, and checking for any problems such as variations in the trigger rate or strange behaviour from individual PMTs.

On a less immediate basis, daily checks are performed considering various physics parameters obtained during the data reduction (to be described later, in Sec. 4.2 and onwards), which generally check for unusual rates of certain classes of data, unusual PMT charge distributions, and so on. On a weekly basis, events that reach the final stages of the reduction (that is, events classified or nearly classified as neutrino events) are manually scanned and investigated, to ensure that the physics abilities of the detector generally remain constant.

3.3 ELECTRONICS AND TRIGGER SYSTEM

Here we discuss both the ID and OD electronics and trigger system. The OD has a similar approach but generally different hardware to the ID, due to the historical fact that it was initially funded and maintained by the American members of the SK collaboration.

3.3.1 *SK-I to SK-III*

The SK-I to SK-III phases used custom built analogue timing modules (ATMs), where each ATM measures the output of 12 PMTs, and converts the analogue charge and timing information to a digital signal. Each PMT channel registers a "hit" if the charge passes a background noise threshold equivalent to $n_{\text{pe}} > 0.25$ (where the exact meaning of

$n_{\rm pe}$ for the SK PMTs will be explained in Sec. 3.4), at which point the time is digitized, and the charge is integrated for the following 400 ns period and then digitized. The digital information is then sent to a central electronics system, and while this is happening the channel is free to be triggered again if necessary.[3] Various "global triggers" are defined, which are issued depending on the number of hit PMTs in a 200 ns time window (which is the approximately the time for Cherenkov light to traverse the detector across the longest possible path). If a global trigger is issued, the integrated charge and time of every PMT hit is saved for a period of $\pm 1.3\,\mu$s around the trigger.

The global triggers initially defined at the beginning of SK-I were the high energy (HE), low energy (LE), and from May 1997 the super low energy (SLE) triggers, which corresponded to requiring 31, 29, and 13 hits respectively,[4] and had trigger rates of approximately 5 Hz, 11 Hz, and 110 Hz. The SLE trigger is defined in order to record solar neutrino events, which have energies $\lesssim 10$ MeV, as 13 hits corresponds roughly to observation of a 4.6 MeV electron (assuming 50% of true hits were detected). From the high frequency it can be seen that the SLE trigger requires extra computational power, and would be heavily dominated by radioactive background events. Work has continued throughout the operation of SK to improve the computer processing techniques, and reduce the background by improving water and air purification techniques, which has allowed the SLE energy threshold to be lowered – the current threshold corresponds to observation of a 3.5 MeV electron, and further tests to lower the threshold are ongoing [81].

The OD trigger operated on the same principle, requiring 19 hits in 200 ns, where again a hit is defined as $n_{\rm pe} > 0.25$ in the OD PMTs. In the OD case however, data is saved for a longer time, from $-10\,\mu$s to $+6\,\mu$s, in order to catch external events such as stopping cosmic ray muons that later produce a Michel electron. An ID trigger will instantly cause an OD trigger, however an OD trigger will cause an ID trigger only after a 100 ns delay (assuming an ID trigger has not occurred naturally in the meantime), which may avoid cutting off information at the end of the ID event.

3.3.2 *SK-IV upgrade*

The SK-IV upgrade replaced both the ID ATM and equivalent OD boards with new custom built components known as QBEEs[5], and a new central electronics system was installed. The QBEEs take 24

---

3 This is achieved by having a pair of converter systems for each channel, such that the secondary system can take over while the primary system is discharging.

4 After subtraction of an average "dark rate", or noise background, of ~ 18 hits per 200 ns.

5 Standing for QTC (charge and time converter) Based Electronics with Ethernet, and are described in [82].

TRIGGER NAME	HIT THRESHOLD	RECORDING PERIOD [μs]
OD	22 (OD)	$-5 \rightarrow +35$
SLE	34	$-0.5 \rightarrow +1$
LE	47	$-5 \rightarrow +35$
HE	50	$-5 \rightarrow +35$
SHE	58	$-5 \rightarrow +35$
AFT	–	$+35 \rightarrow +535$

Table 4: The current trigger types used in SK-IV. "Hit threshold" represents the number of hits required in a 200 ns time window (with no background subtraction). "Recording period" is given relative to the trigger time. The AFT trigger occurs if an SHE trigger has completed without an OD trigger being detected.

inputs each, and have improved charge and timing resolution compared to the ATM boards [79], however the main improvement is the ability to quickly digitize each hit and have them individually reported to a central software-based customizable trigger. In contrast to the inbuilt hardware trigger of SK-I to SK-III, this allows essentially constant monitoring, and when required software recording for longer periods of time. Dead time after initiation of a hardware trigger is also eliminated, so Michel electrons from muon decay can be reconstructed more often.

Different time periods may also be recorded for each type of trigger, as shown in Tab. 4.[6] In particular, low energy events are generally short single-particle events and do not need a large time range, and so the data output may be reduced using a short LE trigger, and a new special high energy (SHE) trigger is also introduced. If a SHE trigger is activated and there is no OD trigger, which is common in contained neutrino events, a further after-trigger (AFT) is issued which records a long time period after the initial event. This is intended to monitor for the γ-ray emission that occurs during neutron capture, which allows for an improvement in the $\nu / \bar{\nu}$ separation [83, 84].[7] Other parameters, such as the n_{pe} threshold and PMT charge integration time, were left unchanged to maintain consistency with the SK-I to SK-III period.

3.3.2.1 *Supernova triggers*

The SK-IV electronics can process up to around 6×10^5 events each second. This is generally enough, even in the case of a supernova neutrino burst within our galaxy, and such an event is monitored for

---

6 in fact, data is recorded in fixed-time blocks such that the recorded data time may be slightly longer than the time windows shown in the table; the recorded data outside the period are however marked as "out of gate" and are generally not used.

7 As anti-neutrino events are more likely to produce neutrons in the final state.

by a dedicated process. Candidate events are reported in real-time on the SuperNova Early Warning System (SNEWS, [85]). Supernovae within $\sim 7\,\text{kpc}$ however, for which there are some possible candidate stars [86], would cause enough neutrino events to overflow the system, and so a dedicated system to record such an event is currently being designed [76].

3.3.3 *High voltage supply*

Each ID PMT requires a high voltage (HV) power supplied, to generate the potential difference over which photoelectrons are accelerated; this was provided by CAEN [87] modules until late 2013, when a new system by iSeg [88] was installed, due to the ageing of the original system. The HV supplies are physically separated into boards, crates, and huts, where in general 24 PMTs are connected to each board, 10 boards are in a crate, and 12 crates are contained within one climate-controlled electronics hut, of which there are four placed on top of the SK detector. The PMT connections to each board generally matches the same set of PMTs connected to a given QBEE (while not particularly interesting from a physics point of view, this structure must be considered when processing data, in the case of electronics downtime or noise events).

3.4 CALIBRATION

3.4.1 *Calibrations of basic detector properties*

Several types of continuous calibration are performed on the detector, to ensure reliable operation even under changes in the water quality, hardware effectiveness, and so on, over time. Both ID and OD PMTs need extensive calibrations in order to understand physically the charge and timing distributions measured by the electronics. Low energy events such as solar neutrinos generally produce only a few PMT hits with a single photoelectron each, for which an accurate knowledge of each PMT's quantum efficiency distribution is paramount, whereas high energy events tend to have multiple photoelectrons per PMT, and the conversion between output charge and number of induced photoelectrons n_{pe} must be well-understood (this is often referred to as the *gain* factor of each PMT).

3.4.1.1 *ID PMT gain calibration*

An example distribution of the charge outputs for a range of n_{pe} was given in Fig. 13, and it is important to note that this kind of distribution can vary on a per-PMT basis. Unfortunately, the electronics used in SK-I→III did not allow discrimination of these distributions on a

per-PMT basis. As such it was necessary to calibrate the detector in the way described below.

The first step in the calibration, most recently performed between SK-II and SK-III, was to determine the correct HV value for 420 reference PMTs. This was done by placing each reference PMT inside a light-proof box with a "scintillator ball" designed to diffuse light uniformly, with an optical input from a wavelength-filtered Xenon flash tube; such a setup allowed an extremely consistent light intensity, and the HV inputs were tuned such that the output charge was the same for each PMT. These reference PMTs were distributed throughout the detector, and the scintillator ball was positioned in the center of the tank. The HV values of the non-reference PMTs were tuned such that their output matched those of the reference PMTs positioned with an equivalent geometrical relationship to the scintillator ball (this method was used as it was previously found difficult to predict the number of n_{pe} expected at a given position mathematically, considering all water quality effects, surface reflections and so on).

After the above procedure, the charge outputs at a particular light intensity and frequency are well-calibrated, however the factor that determines these outputs for each PMT is a combination of both the QE and gain factors. To extract the gain factors, we use a dye laser with variable intensity, injected into a diffuser ball with variable position inside the SK tank. Both high-intensity events (with $n_{\mathrm{pe}} \gg 1$ for every PMT) and low-intensity events (where only a few PMTs are hit) are generated. After many high-intensity flash events, the total charge q_i output by a PMT i is assumed to be given by

$$q_i = I_i \times f_i(\vartheta) \times Q_i \times G_i, \tag{34}$$

where I_i is the totaled intensity of the arriving light, $f_i(\vartheta)$ is a function accounting for angular acceptance differences, Q_i is the QE, and G_i is the gain factor of each PMT. For a low-intensity event however, PMTs are generally hit by a single photon, and the *true* n_{pe} will be 0 or 1 with a probability depending on the QE. The output charge for $n_{\mathrm{pe}} = 1$ should be noticeably above the threshold independently of the gain calibration; that is to say, for a large number of low-intensity flashes the average number of hits n_i is given by

$$n_i = I_i' \times f_i(\vartheta) \times Q_i, \tag{35}$$

where I_i' is the totaled intensity of the arriving light, in the low-intensity case. Comparing the above two equations, we find

$$G_i = \frac{I_i' q_i}{I_i n_i}, \tag{36}$$
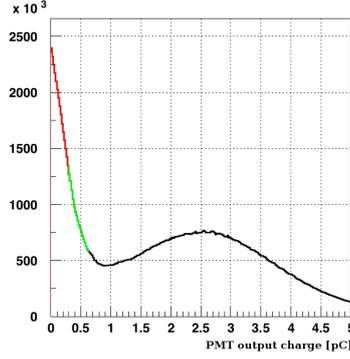
Figure 16: The single photoelectron distribution at SK, using "nickel source" data. The green part of the curve was measured by specially lowering the charge threshold of the ATM modules (see Sec. 3.3), and the red part of the curve is a simple extrapolation. [79]

and assuming I_i'/I_i is roughly constant for each PMT measurement[8] we may define the *relative* gain factor

$$g_i = \frac{q_i}{n_i} \div \left\langle \frac{q}{n} \right\rangle, \tag{37}$$

where the angular brackets denote an average over all PMTs. The relative gain factors g_i are thus the fractional difference in gain from the average, and are distributed around unity, with their 1σ width found to be approximately 7%.

Finally, we can use these relative gain distributions g_i to uncover a more accurate measurement of the average absolute gain $\langle G_i \rangle$ as follows. A "nickel source", which is a radioactive source inside of a nickel ball that causes γ-ray emission after neutron capture, is used to generate a large number of events, > 99% of which are events with single photoelectron emission. The gain corrections $1/g_i$ are applied for each PMT, then the average single photoelectron distribution for every PMT in the ID can be plotted, and is shown in Fig. 16. The fitted value of the single photoelectron peak (the second peak in the figure) then gives the absolute value of the average gain conversion factor, from which individual PMT gains can be measured.

From the above calibrations, it is estimated that the gains and quantum efficiencies of each PMT are known accurately at the 1% level; while this does not directly translate to an equivalent accuracy in the measured light intensities, it allows an accurate Monte Carlo simulation of the detector which shows that we can perform accurate energy reconstructions for contained events from a few MeV up to a few GeV (to be discussed in Sec. 4).

---

8 Which should be true as the light source remains in the same position during each i, and the same power settings are used for the high and low intensity flashes for every i.
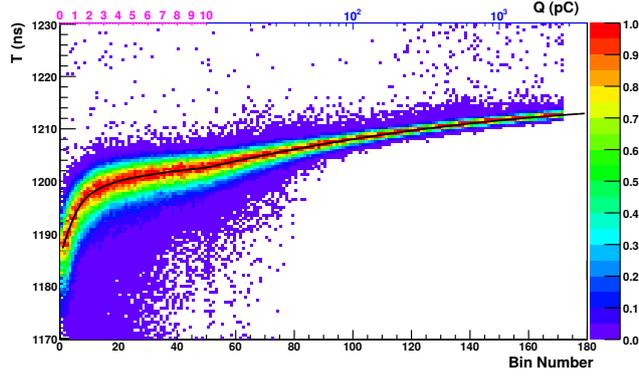
Figure 17: Example time vs. charge distribution map (TQ map) for an SK ID PMT. The upper x-axis showing the actual charge is a combination of linear and log scales, the black line represents a series of fitted polynomial functions, and the color scale is arbitrarily normalized. [79]

3.4.1.2   *ID PMT timing calibration*

The arrival time of the PMT signals must also be calibrated, considering small differences in the physical construction of each PMT, the length of cabling for each PMT, and also the fact that the PMT trigger activation time is itself a function of the total charge; the charge integral will pass the threshold more quickly for a larger current, which is generally correlated with a larger total charge.

The timing calibration again uses a diffuser ball in the center of the tank, of a type that introduces a minimum time spread during diffusion, with light supplied by a nitrogen laser with a short pulse width of ~ 0.4 ns and a variable intensity. The laser output is directly monitored by a fast-response PMT with a time sensitivity of better than 1 ns. Using this system the trigger time of each PMT is recorded as a function of the total charge, as shown in Fig. 17, and a series of polynomial functions are fitted. These functions are used to correct for the variations in timing of each PMT.

After correction, the timing resolution was measured to be on average 2.1 ns at the single photoelectron charge level, improving to 0.5 ns at the ~ 100 photoelectron level. This can be roughly seen to be a satisfactory level, if one considers that Cherenkov light takes on average ~ 100 ns to cross the tank.

3.4.1.3   *ID water properties*

The optical properties of the water are measured using laser light injection from various positions in the detector, and at several wavelengths in the $\lambda = 337 \rightarrow 473$ nm range, which contains the blue part of the visual spectrum where Cherenkov light emission is concentrated. The PMTs directly hit by the laser can be used to measure the amount of light absorbed across the laser's path, while the other

PMTs may measure the amount of light reflected. The intensity of light remaining in the path of the laser $I(x, \lambda)$ at a distance x is expected to follow a simple exponential law

$$I(x, \lambda) = I_0 \exp\left(-\frac{x}{\alpha_{\text{abs}}(\lambda) + \alpha_{\text{sym}}(\lambda) + \alpha_{\text{asym}}(\lambda)}\right) \tag{38}$$

where α_{abs} represents the wavelength-dependent absorption coefficient, and α_{sym} and α_{asym} are the "symmetric" and "asymmetric" scattering coefficients. The "symmetric" coefficient represents Rayleigh scattering and symmetric Mie scattering, and is not actually symmetric but by approximation takes the form $1 + \cos^2 \vartheta$, where ϑ is the angle between the laser direction and the scattered light. The "asymmetric" coefficient represents forward Mie scattering, and takes the simple form $\cos \vartheta$ for $|\vartheta| < 90°$. The parameters are estimated regularly by an automated system, by fitting the laser flash ID events with PMT charge distributions obtained from a laser event Monte Carlo that was generated with various values of the parameters.

3.4.1.4 *OD calibrations*

The OD must also undergo PMT calibrations, although the requirements for charge and timing accuracy are not as stringent as the ID, as the OD data does not generally play as large a role in the fine reconstruction of physics events.[9] In SK-IV, charge calibration is constantly done via comparisons of the dark rate with cosmic ray muon data, where selecting early hits from muon events generally gives single photoelectron triggers, from which the average $n_{\text{pe}} = 1$ charge is directly taken. There is also a fiber-optic system with outputs at various parts of the ID, and the nitrogen laser (used for the ID PMT timing calibration) is also used to provide a set of single photoelectron triggers and give a secondary cross-check. The gain is found to be in the range $1 \to 6$ pC per n_{pe}, and the accuracy of the measured charge was estimated as between 10% and 20%, depending on the particular PMT. The timing offsets are estimated considering the cable length and the laser events, and also by cross-checking timing offsets with the cosmic ray muon data, and after correction the resolution was estimated as better than 10 ns. The water properties of the OD are not measured by direct light injection, but their effects are simulated in a Monte Carlo and tuned to give the best agreement to the physics data.

---

9 For events contained inside the detector, the OD is essentially a veto counter. Even considering uncontained events used in the analysis, such events cannot be accurately reconstructed even in principle, since information is lost once the charged particles are completely outside of the detector.

3.4.1.5   *Other calibrations*

There are several other calibration procedures that provide finer adjustments to the electronics and detector model, however a detailed discussion is left to the published papers [77, 79]. These include accurate geometric measurements, water convection measurements, fine QE and gain measurements by comparison with Monte Carlo simulations, testing non-linearity of charge[10] using a high intensity light source, and measurements of the light reflection properties of the various materials used in the tank.

3.4.2   *Calibrations using physics events*

In principle, from the detector calibrations listed above, we can reconstruct the total intensities of the Cherenkov light observed. Then, while considering the geometry of the detector and physics predictions from Monte Carlo simulation (see Sec. 4.5), we can reconstruct the physics parameters of particle tracks. In order to test the accuracy of these reconstructions (themselves explained in detail in Sec. 4.2.2), and do final calibrations to the reconstructed energy variables, we use several physical processes with known parameters.[11]

3.4.2.1   *Low energy events*

A linear accelerator (or LINAC) is installed above the SK tank, which can produce a beam of electrons of specific energy chosen across the few MeV to tens of MeV range.[12] This is complimented by a neutron generator, which can be lowered into the detector to produce ^{16}N, the decay of which is dominated by an electron and a gamma ray with energies of a few MeV. Using these techniques, the energy resolution after reconstruction was shown to be better than 20% and the vertex resolution within 2 m on an event-by-event basis, even for the absolute lowest energy events. For events with $E_\nu \approx 10\,\text{MeV}$, these parameters improved to $\sim 10\%$ and $< 1\,\text{m}$ respectively.

3.4.2.2   *High energy events*

For higher energies, statistical tests of cosmic ray muon and neutrino-induced pion events are performed. As mentioned previously, the observed energy in photoelectrons is not generally expected to correlate exactly with the energies of the observed particles, and recon-

---

10 In other words, a charge-dependent gain factor in the PMTs or electronics; this is found to occur sometime before the saturation point (highest possible charge observed by the electronics), which is $n_\text{pe} \approx 300$ for SK-I to SK-III and $n_\text{pe} \approx 1200$ for SK-IV.

11 All of these methods are repeated for each of the SK-I to SK-IV periods.

12 While the LINAC was used every few months in early stages of the SK experiment, it was a highly manpower intensive task and was not used often in the later phases of the experiment.

structing the physical parameters of a track requires a comparison with the MC simulation. However, we can at least confirm that the observed charge and reconstructed energy distributions match what is expected by the MC for certain types of high energy events, which gives us confidence in our simulation.

In particular, muons that stop inside of the tank decay to electrons, with an energy given by the well-known Michel spectrum, extending to around 50 MeV. Such events are selected by requiring a stopping muon, identifiable by a clear entering hit cluster in the OD and no corresponding exit cluster, followed by a comparatively low energy second event a few microseconds later.[13] The reconstructed mean energy of these selected events agrees with the mean of the predicted spectrum at the 1% level, and is stable over time.

The cosmic ray muons themselves can also be used; after identifying the muon stopping point using the fitted Michel electron vertex, the track end point can be accurately reconstructed. Furthermore, the expected Cherenkov angle (useful for muons with energy $E_\mu \lesssim 400$ MeV) or the expected energy loss per unit distance (e.g. $\sim 2\,\mathrm{MeV\,cm^{-1}}$ for 1 GeV muons in water) are theoretically well known, and the muon true energy can be quite accurately predicted from the above parameters (denoted as E_{pred}), and compared with the observed total charge q_{obs}. Comparing the above parameters for data events with the equivalent parameters simulated in MC events,

$$\frac{q_{\mathrm{obs}}}{E_{\mathrm{pred}}} \div \frac{q_{\mathrm{obs}}^{\mathrm{MC}}}{E_{\mathrm{pred}}^{\mathrm{MC}}} \tag{39}$$

gives a high-statistics check across a wide range of binned muon energies E_{pred}, and is consistent with unity to within 2%.

A final check comes in the form of the decay of neutral pions, which proceeds by $\pi^0 \rightarrow \gamma + \gamma$. From events with two reconstructed electromagnetic showers fitted (discussed further in Sec. 4.2.2), representing the two γ-rays, the invariant mass of the pion can be reconstructed and compared with the true value of 135 MeV. The position of the mass peak is found to be correct within 2%.

Considering all of the above procedures, across all SK periods and including a wide range of particle energies from tens of MeV up to a few GeV, we estimate the accuracy of the *average* reconstructed event energies to be better than 2%, and the stability of this value over time as better than 1%.[14]

---

13 The fact that the selected events are Michel electrons is confirmed by an accurate fit of the expected decay time of the muon.

14 The event-by-event energy resolution is discussed for various event types in Sec. 4.2.2.

DATA RECONSTRUCTION AND MONTE CARLO

4.1 OVERVIEW

4.1.1 *Event topologies*

The high energy atmospheric neutrino data events are selected and reconstructed into three main samples with different event topologies, in principle containing only neutrino induced events,[1] divided as follows. In the fully-contained (FC) sample, the vertex is inside the fiducial volume and all visible secondary particles are contained inside the inner detector. In the partially-contained (PC) sample the vertex is also inside this fiducial volume, but outgoing particles are allowed to exit the inner detector. PC events have longer charged particle tracks, which are therefore mostly muons induced by ν_μ (or $\bar{\nu}_\mu$) CC interactions. In the upward-going muon event sample (UPMU), neutrinos interacting with the rock create muons which enter the detector from below. Down-going muons are ignored, as these are overwhelmingly produced by cosmic rays. The UPMU sample is also a predominantly ν_μ ($\bar{\nu}_\mu$) induced sample.

The detector in its current configuration has only limited power to distinguish neutrinos from anti-neutrinos. Some separation techniques exist, and will be described and used for one sub-section of the energy spectrum analysis. In most cases however, we discuss and measure only the combined $\nu + \bar{\nu}$ fluxes, therefore further references to ν_μ or ν_e should be taken to imply inclusion of $\bar{\nu}_\mu$ and $\bar{\nu}_e$ respectively, unless explicitly stated.

An overview of the main samples indicating their energy ranges and topologies is shown in Fig. 18 (anticipating some results from later chapters, which provide definitions for dividing each sample into the various sub-samples).

4.1.2 *Reduction and reconstruction*

The recording of high-energy data events occurs based on some simple trigger criteria as defined in Sec. 3.3. The selection of physics events requires first a removal of many background events such as cosmic-ray muons, radioactive decays in the tank, electronics noise

---

1 Assuming the event selections are perfect, and there are no special events such as proton decays. To be clear, this thesis discusses the "high energy" neutrino data, and does not consider the separate processes that are used to select "low energy" (solar) neutrino data, or other specialized samples.
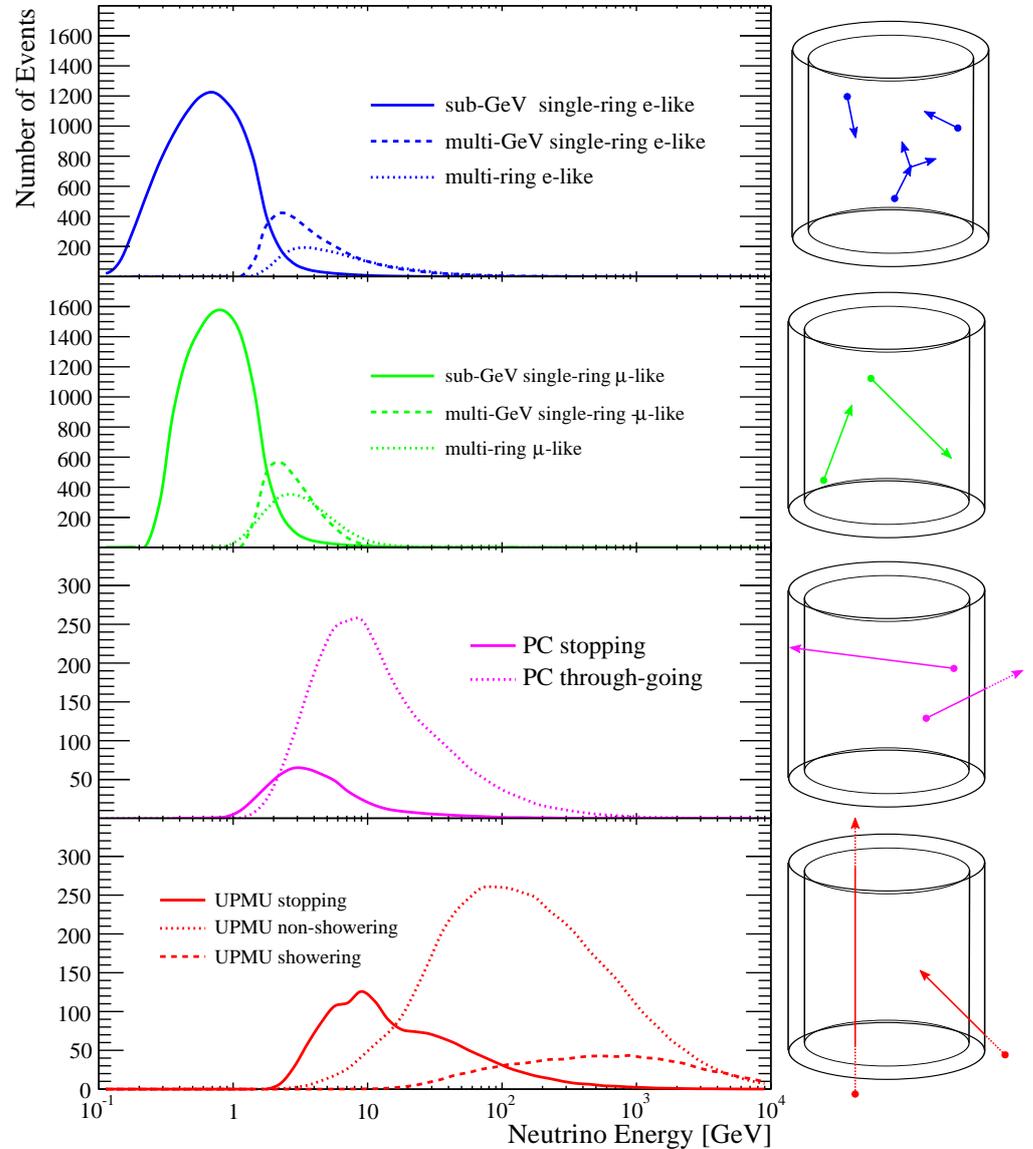
Figure 18: MC neutrino true energy distributions for each sub-sample. From top, the FC *e*-like, FC *μ*-like, PC and UPMU samples, divided into sub-samples as described in Sec. 4.2→Sec. 4.4. Pictorial detector representations show the ID and OD sections, and event topologies.

events, and so on, which is done by a series of increasingly complex cuts and is known as the data *reduction*. Events passing the data reduction have various processes applied to measure the physics parameters of the events, such as the interaction vertex and particle momenta, and is known as the data *reconstruction*.

Each of the three main samples (FC, PC and UPMU) have a distinct set of reduction and reconstruction processes, which will be explained in order below. In general the final sub-sample definitions are evolving, and various SK analyses may use different definitions. Those used in this thesis for flux analysis purposes are similar to those used in the SK atmospheric neutrino oscillation analysis [5], with a few small changes and updates, and without separation of ν and $\bar{\nu}$ events.

4.2 FULLY CONTAINED SAMPLE

4.2.1 *Reduction*[2]

The FC reduction is divided into five steps, denoted as FC1 → FC5, where events that do not pass each step are not processed further in order to optimize the processing time. Several reconstruction programs (or event "fitters") are gradually applied during the reduction, which are mentioned briefly here, and explained further in Sec. 4.2.2 on the FC reconstruction. After the final step, fiducial volume and visible energy cuts are performed.

The reduction cuts were defined by comparing raw data with expectations from the SK atmospheric neutrino MC (Sec. 4.5), aiming to keep at least 99% selection efficiency of *true high-energy neutrino events* for each step, while reducing background to the level of ≪ 1% compared to the neutrino events. Here, MC "true high-energy events" are defined as those with a true vertex inside the FV, with visible energy $E_{\text{vis}} > 30$ MeV (which is a parameter that will be defined in Sec. 4.2.2), and with less than 16 clustered hits in the OD (less than 10, in the case of SK-II).

Some short notes on reading the below cuts: many are given in terms of estimated number of photoelectrons (n_{pe}), which are converted from PMT charge outputs as explained in Sec. 3.4, where tubes that are not "hit" do not contribute to n_{pe} totals; secondly, the trigger timing is explained in Sec. 3.3, and if cuts do not specify a time range, they should be assumed to cover the standard range of $\pm 1.3\,\mu s$; finally, all cuts are given here in terms of the requirements for *passing* each cut.

---

2 The author of this thesis was in charge of the FC data reduction process, and so this section goes into extensive details (and was thoroughly checked) for the purposes of future reference.

4.2.1.1   *FC1*

The first reduction implements very basic cuts:

1. total ID n_{pe} in a sliding 300 ns time window > 200 (for SK-II, > 100)

which allows only events with significant ID activity, cutting low energy events such as radioactive background and solar neutrinos. The SK-II criteria is relaxed, in this and many other cuts, considering that number of active PMTs was approximately half that of the other periods. Then,

2.   a) number of OD tubes hit > 0

   b) *and* number of OD tubes hit within a time window -500 ns to $+300$ ns is ≤ 50 (for SK-IV, ≤ 55)

which avoids events in the rare case of no OD information (i.e. OD dead-time), and events with obvious OD activity, such as cosmic ray muons (entering before the trigger) and PC events (exiting after the trigger). The criteria is slightly higher in SK-IV, due to the replacement of the OD electronics boards and changes in the PMT gains, and the cut is tuned to reject similar muon energies as the earlier periods.

After FC1, the event rate is reduced from $\sim 10^6$ events per day to approximately 2,500 events per day. The efficiency for selecting true high-energy neutrino events is estimated by applying the reduction to the atmospheric neutrino MC and is found to be 100.0% for FC1.[3]

4.2.1.2   *FC2*

The second reduction continues with more simple cuts:

3. highest n_{pe} for an ID PMT divided by the total ID $n_{\mathrm{pe}} < 0.5$

which removes electrical noise events whereby the majority of the measured charge comes from a single tube, then

4.   a) number of OD tubes hit within a time window -500 ns to $+300$ ns is ≤ 25 (for SK-IV, ≤ 30)

   b) *or* the total n_{pe} of the ID is $>100,000$ (for SK-II, $> 50,000$)

which avoids events with OD activity in a stricter sense than cut 2b, while allowing extremely energetic events in the ID, which are expected to leak some charge into the OD even if the particle is fully-contained (e.g., by electrical cross-talk between channels). The effects of cuts 3 and 4a are illustrated graphically in Fig. 19, showing the atmospheric MC and raw data distributions.

---

3 The event rates and efficiencies given in this section refer to the SK-IV period, but other periods are similar.
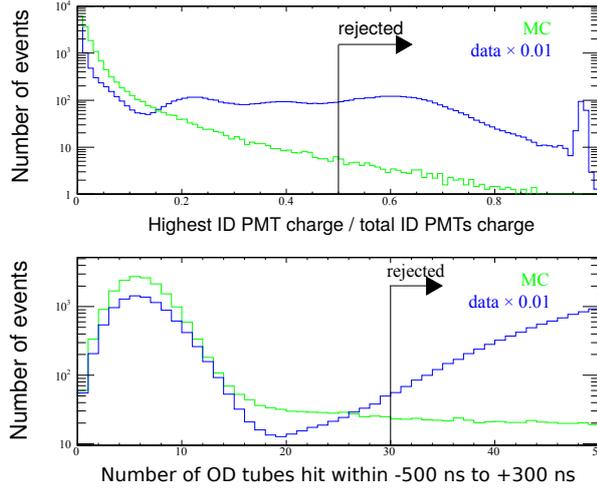
Figure 19: Distributions of the atmospheric MC and data before the FC2 cuts 3 (top) and 4a (bottom), where MC events are true high-energy FC neutrino events in the FV, and data represents the raw detector output before each cut. Note the logarithmic y-axis and the scaling factor of the data distribution. [89]

After FC2, the event rate is reduced to approximately 500 events per day, and the efficiency of FC2 to select true neutrino events is again estimated as 100.0%.

4.2.1.3   FC3

The third reduction features more complicated cuts, which are designed to pick out particular types of background events. Several of these cuts target cosmic-ray muons; the first cut targets "hard" muons, which are muons with $E_\mu \gtrsim 1\,\mathrm{TeV}$ whose energy loss by processes such as pair production and bremsstrahlung is greater than their ionization energy losses, and can thus cause many hits in a short time:

5. number of OD tubes hit within a sliding 500 ns time window < 40.

The next cut targets through-going muons, by applying a simple fitter that can find two clusters of hits in the OD, corresponding to an entering and exiting muon with an appropriate travel time. The fitter is only applied if the ID contains more than 231 photoelectrons spread over 1,000 hit tubes, however once the fitter is applied events may still pass as follows:

6.   a) through-going muon goodness of fit < 0.75

b) *or* number of hits in both the entrance and exit OD hit clusters < 10.

A similar fit for stopping muons is also applied, which finds an entering but not an exiting cluster, and events may pass by having
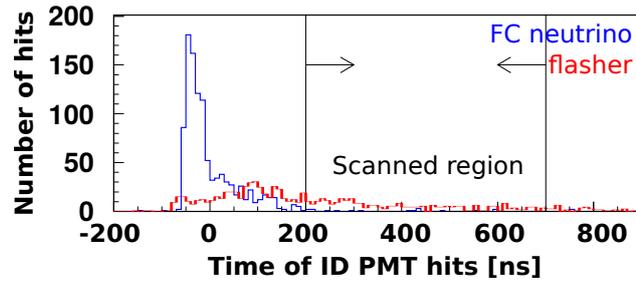
Figure 20: Comparison of a nominal FC neutrino event and a nominal flasher event, both found in SK-I data. The scanned region shows the timing region tested by the FC3 cut 10. [90]

7. a) stopping-muon goodness of fit ≤ 0.0

   b) *or* number of hits in the entrance OD hit cluster < 10,

and for SK-I only, a further finely-tuned cut is applied as

8. a) stopping-muon goodness of fit < 0.5

   b) *or* number of hits in the entrance OD hit cluster < 5.

The next cut targets remaining low energy events, using a simple time-of-flight (TOF) fitter "point-fit" to obtain an event vertex based on the timing information of each PMT and the speed of light in water. Assuming all photons were generated at this vertex, the TOF of each hit is subtracted, and events must pass

9. number of time-of-flight corrected hits in the ID within a sliding 50 ns time window ≥ 50 (for SK-II, ≥ 25).

The next cuts remove so-called "flasher" events, which are events whereby electrical discharge inside a PMT causes some emission of light, which may look like a contained event. In general the timing distribution of hits is longer for flashers compared to real particle track events; an example is shown in Fig. 20. For SK-II to SK-IV a cut is thus made as

10. minimum number of hits in the ID within a sliding 100 ns time window between $+200$ ns and $+700$ ns is < 20

which disallows flasher events with a long tail in the hit timing. The SK-I cut is slightly more complicated, as it was tuned in a tighter sense. For SK-I only,

11. a) minimum number of hits in the ID within a sliding 100 ns time window between $+200$ ns and $+700$ ns is < 10

   b) *or*

      i. the above count is < 15

      ii. *and* total number of hits in the ID is > 800

which essentially places a tighter constraint, unless the event was particularly energetic, then a medium constraint is placed. Low-energy flashers are also cut by a goodness-of-fit cut on the point-fit TOF fitter as

12. TOF fitter goodness-of-fit > 0.4

since for wide time distributions, a simple time-of-flight vertex fit searching for a single point will not work well. The next event type targeted are the "coincidence muon" events, which are due to the high frequency of SLE triggers, during which cosmic-ray muons sometimes enter the detector by coincidence. This requires either no late OD activity, or an ID activity lower than most cosmic-ray muons, as

13.   a) number of OD hits between +300 ns and +800 ns is < 20

   b) *or* total ID n_{pe} count is $\leq 5,000$ (for SK-II, $< 2,500$).

The final third reduction cut removes "cable hole muons". The cabling for the PMTs enters the SK superstructure through twelve holes in the top of the tank, and (due to the need to insulate the high-voltage power supply) the radius of the cabling bundle is quite wide at around 1 m. Down-going muons could theoretically pass through the cable holes and be reconstructed as FC events; to test for such events, four of the cable holes are covered with plastic scintillator veto counters. In conjunction with the stopping-muon fitter, the cut

14.   a) number of veto-counter hits $= 0$

   b) *or* stopping-muon best fit entry-point is > 4 m from the cable holes (in both x and y directions)

is applied, which ends the third reduction.

   After FC3, the event rate has been significantly reduced to around 50 events per day, and all events passing the third reduction have the "APFIT" software applied; this is the complex high-energy fitter and applies the fitting processes described in Sec. 4.2.2. The efficiency of FC3 to select true neutrino events is estimated to be 99.8%.

4.2.1.4   *FC4*

The fourth reduction specifically targets flasher events, that were not removed by the above cuts, by use of a pattern-matching algorithm. Such an algorithm is useful because flashing tubes often emit light in a specific way, and create similar hit patterns within the tank repeatedly. The algorithm divides the detector into many $4\,\mathrm{m}^2$ patches, sums the n_{pe} in each patch, and compares the similarity of the pattern to all other events in that SK period by the parameter r, which is defined between two events A and B as

$$r = \frac{1}{N} \sum_i^N \frac{\left(Q_i^A - \langle Q^A \rangle\right) \times \left(Q_i^B - \langle Q^B \rangle\right)}{\sigma^A \sigma^B} \tag{40}$$
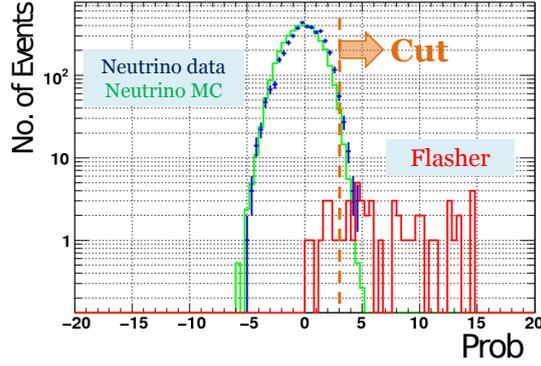
Figure 21: Comparison of the likelihood parameter "Prob" for neutrino data, MC, and hand-selected flasher events, showing the cut value. [76]

where N is the number of patches, Q_i is the summed n_{pe} in each patch i, and σ is the RMS of the Q_i.

A further parameter d_{KS} is also introduced, comparing the time distribution of the charge by a Kolmogorov-Smirnov test,[4] and a third parameter is defined depending on whether or not the PMT with the highest n_{pe} (which is often the flashing tube) matches between events. These parameters are combined into a single likelihood parameter "Prob", and if two events score higher than a cut value (determined based on a comparison of Monte Carlo neutrinos against hand-selected real flasher events), they are both removed as flasher events, as shown in Fig. 21.

The fourth reduction removes on average only a few events per day, with a neutrino efficiency of 99.6%.

4.2.1.5   FC5

The fifth reduction again targets several specific background types with tailored cuts, generally targeting muons in a stricter sense than the third reduction, along with other cuts on particular types of electronic noise events and calibration-related events. The first cut targets "invisible muons", which are cosmic-ray muons below the Cherenkov threshold that are undetectable, but stop in the tank and decay via the Michel decay. This can create an electron above its Cherenkov threshold originating inside the fiducial volume, which could be mistaken as an FC event. These events are removed by requiring

15.   a) total ID $n_{\text{pe}} > 1{,}000$ (for SK-II, > 500)

   b) *or* number of OD tubes hit within a sliding 200 ns time window from -9000 ns to -200 ns is < 5

---

4 The "K-S" test uses as a statistic the largest difference in the integrated charge distributions over time, and is essentially sensitive to shifts in position or shape of the time distributions of hit PMTs.

c) *or* the above count plus the number of OD tubes hit within -200 ns to +300 ns is < 10 (applied only if the distance between the hit clusters in the two time periods is within 0.5 m)

where the first cut allows all events above the maximum energy of the Michel decay,[5] and the next two cuts allow events below this energy only if they have insignificant early OD activity.[6] The second fifth-reduction cut targets again coincidence muons, and requires

16. a) the number of OD tubes hit within a sliding 200 ns time window from +300 ns to +2100 ns is < 20

b) *or* total ID $n_{\mathrm{pe}} > 300$ (for SK-II, > 150)

which scans over a longer time range than cut 13a, but allows events if they were not caused by SLE triggers (and thus statistically unlikely to produce coincidence events). Several cuts to remove stopping muons, in a tighter sense than previous cuts, are then performed. A clustering algorithm is run on the OD data, identifying OD tubes that are hit close together, and the following cut is performed for SK-I only:

17. number of OD PMTs found in a cluster < 10.

The third reduction cut 8a is then applied on SK-II to SK-IV data at this stage. Furthermore, a tighter cut on cable-hole-muons focusing on the cable holes which do not have veto counters is performed for SK-IV only:

18. a) stopping muon fitter goodness-of-fit < 0.4

b) *or* total ID $n_{\mathrm{pe}} < 4000$

c) *or* the fitted muon's directional unit vector has $\cos(z) > -0.6$

d) *or* the distance between the stopping muon's best-fit entry-point and the closest cable hole is >2.5 m

and a cut based on APFIT fitted muon rings is introduced, where

19. number of OD tubes hit within 8 m from the APFIT fitted entry point within a sliding 200 ns time window from -500 ns to $+300$ ns is < 4.

The next cut is another flasher cut, targeting wide hit-timing distributions in a stricter sense than cuts 10 and 11. However, this cut also uses a more advanced TOF-fitter, and allows events if the timing distribution corresponds well to the fitted particle tracks:

---

5 The parent muon has almost certainly come to rest in the lab frame.

6 Even "invisible muons" would be expected to have some OD activity, as if they could reach the fiducial volume, they would have had to have been above the Cherenkov threshold during passage through the OD.

20. a) the minimum number of ID tubes hit within a sliding 100 ns time window from +200 ns to +700 ns is < 6

b) *or* TOF fitter goodness-of-fit > 0.4 ,

then and a stricter criteria for the TOF fitter is applied on events with a wide hit-timing distribution (applied for SK-II to SK-IV only):

21. a) the minimum number of ID tubes hit within a sliding 100 ns time window from +200 ns to +700 ns is ≥ 6

b) *and* TOF fitter goodness-of-fit > 0.3.

The rest of the cuts remove fake events occurring due to various particularities of the SK electronics (and so are less interesting from a physics perspective, but nonetheless important). The first such cut targets electronic noise events coming from signal leakage in the high-voltage systems or electronics boards, which tend to result in a large number of fake tube hits at low n_{pe} values. We require

22. a) number of ID tubes hit with less than a single photoelectron is < 250 (for SK-II, < 125)

b) *or* the above count minus the number of ID tubes hit with greater than a single photoelectron is < 100 (for SK-II, < 50).

The next cut removes events where any of the electronics huts (each accounting for one quarter of the detector) were temporarily off for that event.[7] A dead hut can be assumed if all tubes associated with the hut report no hits.

23. Number of dead electronics huts = 0.

The next cut targets events where the OD timing is out of sync with the ID timing, which was a problem identified in the early stages of the experiment in some high-energy events, and is applied for SK-I to SK-III only. It can be identified by large early-timing clusters, or the trigger counting an unusual number of sub-events:

24. a) total ID n_{pe} < 1000 (for SK-II, < 500)

b) *or*

 i. number of OD tubes hit within a sliding 200 ns time window from −9000 ns to −200 ns is < 20

 ii. *or* number of sub events ! = 7.

The final two cuts remove laser and xenon calibration events (see Sec. 3.4). Although these events should be tagged and removed at calibration time, some events remain in the data, and are thus removed based on signals from dedicated PMTs directly monitoring the calibration sources:

7 Monte Carlo studies showed that smaller problems, such as deadtime for individual boards, do not strongly compromise the ability of the detector to correctly reconstruct events.

25. laser and Xenon monitor PMT hits $= 0$

which ends the fifth reduction.

Again, only a few events per day are removed by the fifth reduction, and the estimated neutrino efficiency is 99.9%.

4.2.1.6   *Final Reduction*

Finally, although not directly applied in the reduction software, the following cuts are required for events to be classified into the high-energy dataset:

26. the distance from the neutrino interaction best-fit vertex and the wall is $> 2\,\mathrm{m}$,

27. the number of OD PMTs found by a clustering algorithm is < 16 (for SK-II, < 10),

28. the visible energy $E_{vis} > 30\,\mathrm{MeV}$,

where the fiducial volume cut 26 and the visible energy cut 28 (where E_{vis} will be defined in Sec. 4.2.2, and corresponds roughly to the energy scale of the event) are expected to remove the majority of remaining low-energy background events, such as radioactivities.

After the above cuts, the event rate is approximately 8 events per day. The neutrino efficiency of these cuts is estimated to be 99.2%.

4.2.1.7   *Summary*

After all of the above steps, the final FC dataset has daily event rates as shown in Tab. 5, for data and MC for SK-I to SK-IV. We see that the predicted MC rates are slightly lower than the data rates – however these numbers are generated using a simplified two-flavour oscillation scheme, and do not take any systematic errors into account. In the final oscillation analysis, the agreement is marginally improved after applying three-flavour oscillation, and the data rates are within the estimated systematic error range. [5]

The efficiency in the table is estimated (as described above) using the atmospheric MC, by the surviving percentage of FC high-energy events inside the FV. The uncertainties on these numbers were estimated by a study [89] where each cut parameter was varied slightly; these uncertainties are later considered as a systematic error. The background, on the other hand, is estimated by direct eye-scanning of all events passing FC5, during which events such as flashers and cosmic-ray muons are tagged (but *not* removed from the analysis, in order to avoid a bias in the data and MC selection processes). In the table, we show the percentage of such tagged events in the dataset

	EVENT RATE $[\text{DAY}^{-1}]$		EFFICIENCY [%]	BACKGROUND [%]
	DATA	MC		
SK-I	8.26 ± 0.06	7.75	97.75 ± 0.20	–
SK-II	8.24 ± 0.10	7.73	98.31 ± 0.20	0.30
SK-III	8.41 ± 0.13	7.91	97.96 ± 0.80	0.39
SK-IV	8.12 ± 0.06	7.94	98.47 ± 0.30	0.02

Table 5: The final daily event rates of events passing the FC reduction, with statistical error (negligible for MC). The meanings of "efficiency" and "background" are explained in the text.

after the final reduction.[8] No eye-scanning data is available for SK-I, but the background rate is expected to be similar to the other periods.

4.2.2 *Reconstruction*

The job of the FC reconstruction is to estimate the neutrino's flavour and momentum as closely as possible. Fully-contained events can however range from simple single-lepton-track events, to complicated multi-track hadronic events. As the accuracy of the neutrino reconstruction depends on the type of event, sub-samples are defined which group together similar events. For example, single-ring events have relatively higher purity in neutrino flavor, while multi-ring events tend to cover a higher neutrino energy range.

This section will describe the various processes used to assign event properties, and then the division of events into final sub-samples. While the important principles of the fitting procedures are discussed, more extensive technical details can be found in [91]. Furthermore, while various final state products of neutrino interactions are mentioned here, a detailed discussion of the neutrino interactions which create them will be given later in Sec. 4.5.1.

4.2.2.1 *Vertex fitting*

The neutrino interaction vertex is reconstructed in a multiple step process: firstly, a "point-fit" time-of-flight algorithm roughly estimates the vertex position, then the edge of the primary Cherenkov ring emanating from that vertex is fitted, after which corrections due to finite particle track length are precisely fitted and the vertex is iteratively improved. Thus, the vertex fitting is closely interrelated with the primary ring fitting.

The point-fit algorithm assumes that Cherenkov light is emitted from a single point inside the detector at an unknown time. Taking

---

8 Depending on the particular reconstruction and selection procedure of a given analysis, sometimes more conservative systematic errors on the backgrounds are defined.
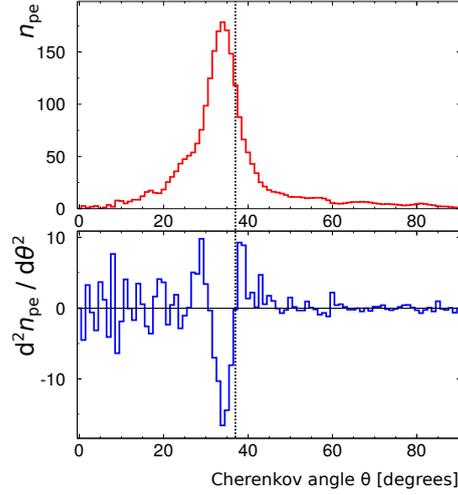
Figure 22: An example distribution of observed charge depending on the Cherenkov angle θ, and its second derivative, for a given ring position and direction. The dotted line indicates the fitted edge of the ring. [66]

an initial test point $\mathbf{x}$, the time-of-flight between that point and each hit PMT can be calculated, and by subtracting the expected time for light to travel those distances we define a set of "residual timings" $t_i(\mathbf{x})$ for each hit PMT i. If the observed light in the event was indeed generated at that point, the $t_i(\mathbf{x})$ should be narrowly distributed at the offset t_0 (the interaction time), with some Gaussian smearing given by the PMT timing resolution. The Gaussian parameter

$$G_p = \sum_i e^{-(t_i(\mathbf{x})-t_0)^2} \tag{41}$$

is largest when the distribution is narrow, so maximising G simultaneously for the point $\mathbf{x}$ and time t_0 gives the "point-fit vertex".

The second step is to identify the primary Cherenkov ring, which should be the highest charge ring in a multi-ring event. The direction of the charged particles leaving the vertex is established in a general sense by defining a vector from the point-fit vertex to each hit PMT, weighted by the observed charge at that PMT, and summing all such vectors. Using this vector as an ansatz, we can plot the charge distribution $n_{\mathrm{pe}}(\theta)$ as a function of the Cherenkov angle to the ansatz vector θ, making careful corrections for the geometry of the detector, the light attenuation in water, and the angular acceptance function of the PMTs. An example distribution is shown in Fig. 22. To find the outer edge of the ring θ_{edge}, which is to say the estimated initial Cherenkov angle, we find the first turning point

$$\frac{d^2 n_{\mathrm{pe}}(\theta)}{d\theta^2} = 0 \tag{42}$$

that occurs *after* the peak value of $n_{pe}(\theta)$, as illustrated in Fig. 22. A goodness-of-fit parameter $Q(\theta_{edge})$ for the fitted ring is defined as

$$Q(\theta_{edge}) = \frac{\int_0^{\theta_{edge}} n_{pe}(\theta) d\theta}{\sin \theta_{edge}} \times \left(\left[\frac{dn_{pe}(\theta)}{d\theta} \right]_{\theta=\theta_{edge}} \right)^2 \times \exp\left(-\frac{(\theta_{edge} - \theta_{exp})^2}{2\sigma^2} \right)$$

(43)

where θ_{exp} is the theoretically expected Cherenkov angle based on the total charge in the cone, and σ is the expected resolution of the edge search (based on the distance between ID PMTs). This parameter is essentially designed to prefer fitted cones containing a large number of photoelectrons (first term), with a sharp edge (second term), and a fitted edge near to the expected value (third term). To penalize unphysical Cherenkov cones, in the case of $\theta_{edge} > 43°$ the second term is removed. The parameter $Q(\theta_{edge})$ is then maximised by shifting θ_{edge}, and also allowing shifts in the ansatz vector, which may improve the fitting. The PMTs intersecting the edge of the best-fit ring are tagged for use in the next step.

Final adjustments are performed by the fitter TDCFit, which considers a finite track length for the primary Cherenkov particle, and finds the best-fit start and end points. The track length is initially estimated by counting the observed n_{pe} and assuming the energy deposited per unit length is $dE/dx = 3\,\text{MeV cm}^{-1}$ (which is at least a reasonable value for electrons or muons of some energies, and was empirically found to result in good fits for many cases without getting stuck in local minima). New timing residuals are calculated based on emission of photons at the same Cherenkov angle along the particle track, and furthermore the PMTs outside of the tagged ring edge are also considered, taking into account the scattering parameters estimated during water quality checks. The track length, vertex position, and ring direction are then iteratively tuned until a local minimum is found. The performance of the vertex fitting procedure is tested with the atmospheric MC events, and the resolution[9] of the vertex fit at this stage is found to be better than 0.7 m in position, and a few degrees in direction, even for complicated multi-ring events.

Finally, in events where secondary rings are not found, and the flavour of the lepton is established (as to be described in the next sections) a further precise ring fitting procedure called "MS fit" is applied. Various parameters, such as dE/dx, can be better predicted depending on the lepton type, and a more precise fitting is possible. After the improvements by MS fit, the resolution for single-ring events is found to be better than ~ 0.3 m in position and $\sim 3°$ in direction.

---

9 Defined as the distance from the true vertex within which 68% of fitted vertices are found.

(a) Image space   (b) Hough space
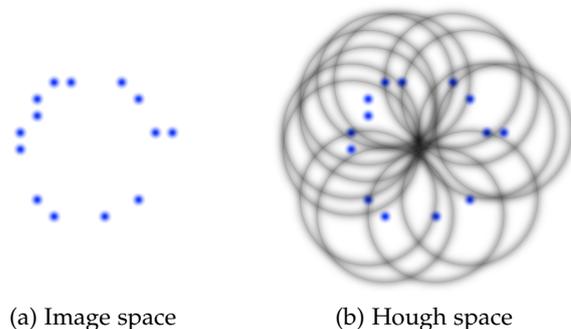
Figure 23: Illustration of the principle of the circular Hough transform, where blue represents image intensity, and black represents the Hough parameter intensity. The recorded image is shown in (a), and the transformation to circular Hough space in (b), where the location of the darkest point is the estimated centre of the fixed width circle (the image itself is overlaid in the Hough space, for illustrative purposes).

4.2.2.2   *Ring counting*

After the vertex fitting, it is expected that the primary (i.e. brightest) Cherenkov ring has been well fitted; the next step is to search for possible secondary rings.

A common class of method of searching for a given pattern or shape in an image is known as the Hough transform; by moving to a new parameter space based on the original image, for example intensity or color as a function of some different co-ordinate system, features of the image may be extracted. The circular case is particularly simple, as illustrated in Fig. 23, where the Hough space can be thought of as a convolution of the original image with a circular function with the expected radius, and has the same co-ordinate space as the original image. The dark points in the Hough space are likely locations for the center of a circle.

The SK case follows essentially the same principle, with of course corrections for the geometry of the detector and the behavior of Cherenkov light in water. The observed n_{pe} fitted as belonging to the primary ring are also subtracted, to search for new rings only, originating from the previously fitted vertex. The initial convolution pattern used is the expected Cherenkov distribution for a 500 MeV electron. An example Hough search at SK is shown in Fig. 24, where the peaks represent likely Cherenkov rings. Peaks found close to the primary ring however (within 15°) are discarded as remnants of the primary ring, which may have come from e.g. a kink in the particle track.

Once candidates are identified from the Hough peaks, they are each tested by a likelihood method to decide if they are true rings or not. The likelihood $L(N)$ for testing each new peak (in addition to the $N-1$ peaks already fitted) essentially tests the observed charge distri-
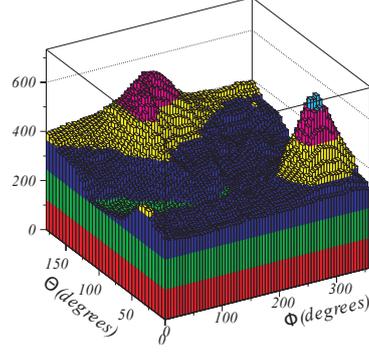
Figure 24: An example Hough-space for a ring search in SK, showing the value of a Hough parameter depending on the zenith Θ and azimuth ϕ angles from the fitted vertex, where two rings are clearly found (in this figure, the n_{pe} of the primary ring is not subtracted). [66]

bution $n_{\text{pe},i}^{\text{obs}}$ for each PMT i against the expected distribution $n_{\text{pe},i}^{exp}(\boldsymbol{\eta})$, which is a distribution that can be varied by changing the estimated physics parameters $\boldsymbol{\eta}$ (such as particle type and energy) of the new Nth ring. The likelihood is thus written

$$L(N) = \sum_i \ln P\left(n_{\text{pe},i}^{\text{obs}} | n_{\text{pe},i}^{\text{exp}}(\boldsymbol{\eta})\right) \tag{44}$$

where the probability function P comparing the charge of each PMT is defined as a Poisson distribution convoluted with the single-photoelectron distribution for $n_{\text{pe},i}^{\text{exp}}(\boldsymbol{\eta}) < 20$, and simplified to a Gaussian for $n_{\text{pe},i}^{\text{exp}}(\boldsymbol{\eta}) \geq 20$. The maximal $L(N)$ is determined by finding the best-fit $n_{\text{pe},i}^{\text{exp}}(\boldsymbol{\eta})$. Due to the fact that extra peaks must represent at least some energy measured by the PMTs, there is always some charge that must be accounted for, which means the momentum physics parameter in $\boldsymbol{\eta}$ has some minimum value. If the particle track is well fitted then $L(N) > L(N-1)$, however due to the minimum momentum requirement a bad fitting will cause a decrease in $L(N)$, so we may have $L(N) < L(N-1)$.

The final choice of whether or not a ring is fitted as true is not actually directly based on the $L(N) - L(N-1)$ parameter, but by a more detailed processing of the two charge distributions obtained in the $L(N)$ and $L(N-1)$ cases, summarized in a final obtained parameter denoted L'.[10] Extra rings are searched for, until none can be found that increase the best-fit likelihood L'. The final parameter $L'(N) - L'(N=1)$, which denotes the single-ring vs. multi-ring likelihood, is shown for SK-IV in Fig. 25, divided into two samples based on the energy. It can be seen that true CCQE events, in which the initial neutrino interactions are free from complicated hadronic effects,

---

10 L' uses a weighted $L(N) - L(N-1)$ combined with extra shape fitting of the charge distributions and timing corrections; details in [91].

(a)   (b)
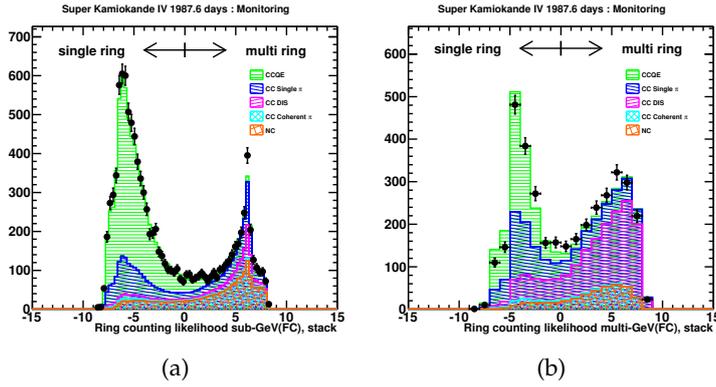
Figure 25: Ring-counting likelihood separations for the sub-GeV and multi-GeV samples for SK-IV, showing the number of MC (normalized stack plot by true interaction type) and the data (points with statistical error bars) event counts. In this and other figures in this section, the MC is shown without systematic error pulls and using a simple two-flavour oscillation model.

(a) *e*-like   (b) *μ*-like
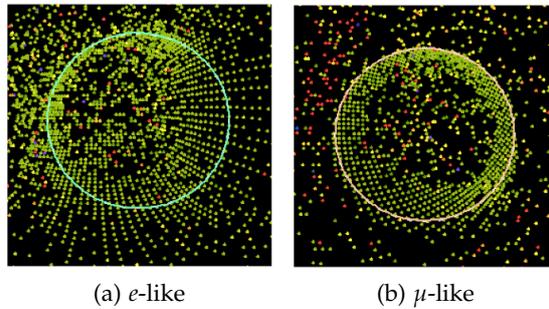
Figure 26: An example comparison between an *e*-like and a *μ*-like ring pattern of roughly the same energy, where each dot represents a hit PMT, colored corresponding to the detection time. [66]

are often identified as single-ring events (approximately 90% of the time) as expected.

Multi-ring events undergo a further processing to cleanly separate and assign the n_{pe} contributions of overlapping rings. Also, after the particle flavour identification has been later performed, the rings are re-fitted using this new information, which can result in further corrections to the final number of fitted rings.

4.2.2.3 *Particle identification*

The detector has excellent particle identification (PID) capability by using the Cherenkov ring pattern and opening angle. Electrons or high energy photons undergo electromagnetic showering and multiple scattering, and create fuzzy Cherenkov rings; in contrast meson tracks, such as muons or pions, are generally straight and create Cherenkov rings with sharply-defined edges, as shown in Fig. 26.
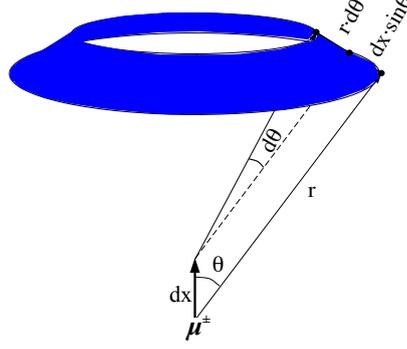
Figure 27: Diagram showing the calculation of Cherenkov light emission from a muon track (variables are explained in the text). [66]

The electron and muon hypotheses are compared for each ring using a likelihood parameter, which takes as input the difference between the measured and expected charge distributions (as in the case of ring counting). For electrons, the expected charge distributions $n^e_{\text{pe},i}$ are obtained by interpolating tabulated distributions at various fixed energies, taken from high-statistics SK MC simulations. For muons however, the expected distribution can be obtained analytically each time. Essentially the muon is propagated a distance x from the initial vertex, during which the energy losses cause a decrease in the Cherenkov angle θ, as shown in Fig. 27 for a propagated distance dx. Each PMT then receives photons from roughly a single position of the muon track, and the expected charge observed at each PMT i is given (to first order, without normalization) by

$$
n^{\mu}_{\text{pe},i} = \frac{\sin^2 \theta_i}{|\mathbf{r}_i| \left(\sin \theta_i + \mathbf{r}_i \cdot [\partial_x \theta]_{x=x_i} \right)} \, e^{-|\mathbf{r}_i|/\alpha} \, f_i(\vartheta) \tag{45}
$$

where x_i is the distance along the track to the point from which the PMT i receives Cherenkov light, $\mathbf{r}_i$ is the vector from that point to the PMT, and θ_i is the Cherenkov angle at that point. In the first term, the numerator is a factor accounting for the dependence of emitted light intensity as a function of the Cherenkov angle, and the denominator accounts for the light intensity at the PMT i, as can be calculated by considering Fig. 27. The second term controls the attenuation of light in water, where α is the attenuation coefficient, and the third term $f_i(\theta)$ is the angular acceptance function of the PMTs. In fact, extra terms are also included to account for scattered light and knock-on electrons created in collisions, but we do not discuss these terms in detail here.

The expected distributions $n^e_{\text{pe},i}$ and $n^{\mu}_{\text{pe},i}$ are both fit the observed distribution $n^{\text{obs}}_{\text{pe},i}$ as closely as possible by optimizing the physics parameters, such as the track direction and the initial Cherenkov angle.

The likelihood parameter L_{PID} of each ring is then calculated using these best-fit distributions as

$$L_{\text{PID}} = \sum_i \ln \frac{P\left(n^{\text{obs}}_{\text{pe},i} | n^{\mu}_{\text{pe},i}\right)}{P\left(n^{\text{obs}}_{\text{pe},i} | n^{e}_{\text{pe},i}\right)} \tag{46}$$

where, similarly to Eq. (44), the distribution P is defined as a Poisson distribution convoluted with the single-photoelectron distribution for $n^{\text{obs}}_{\text{pe},i} < 20$, and a Gaussian for $n^{\text{obs}}_{\text{pe},i} \geq 20$.

For the single-ring case, events are denoted as e-like or μ-like depending on the type with the highest likelihood. Figure 28 shows the L_{PID} separation overlaid with MC truth information, from which we can see that the e-like and μ-like samples correspond very well to ν_e and ν_μ CCQE interactions respectively, with a tiny leakage between each flavour at the $< 1\%$ level. Some backgrounds however remain whereby the flavour is difficult to reconstruct, including neutral current and single pion events, at the few percent level for the sub-GeV and somewhat worse for the multi-GeV events.

Multi-ring events can be coarsely separated based on the PID of the most energetic ring, however depending on the energy range and interaction type, this separation is generally not as good as the single-ring case (especially for sub-GeV events), with wrong-flavour background at the 30% level. We thus use a further procedure, as described in [5] but extended to include SK-IV, to produce a reduced multi-ring sample that can be separated into e-like or μ-like categories more accurately. Using the energy reconstruction definitions to be explained in the next subsection, events with $E_{\text{vis}} > 600$ MeV, and where the primary ring is μ-like with $E_{\text{rec}} > 600$ MeV, are selected as μ-like. To select e-like events, we require $E_{\text{vis}} > 1330$ MeV, and the event to pass a likelihood-based selection that considers the type, momentum, and timing of all fitted rings. This reduces background contamination from hadronic showers, produced by NC or ν_μ CC events, which are often dominated by π^0 and mis-reconstructed as e-like. These multi-ring samples have a flavour purity of 96.6% for μ-like and 77.8% for e-like, in SK-IV.

4.2.2.4   *Energy reconstruction*

From the above vertex and ring fitting processes, we have established for each Cherenkov ring a PID and a tagged list of PMTs (and the fraction of n_{pe} registered in each PMT which is estimated to belong to that ring specifically). The reconstruction of the momentum of the lepton that created the ring then uses a different method depending on whether the lepton is e-like or μ-like. In the electron case, the momentum is taken as being correlated with a simple count of the n_{pe} associated with the ring, whereas for a muon, measurement of

(a) True CCQE electrons   (b) True CCQE muons   (c) All interaction types

(d) True CCQE electrons   (e) True CCQE muons   (f) All interaction types
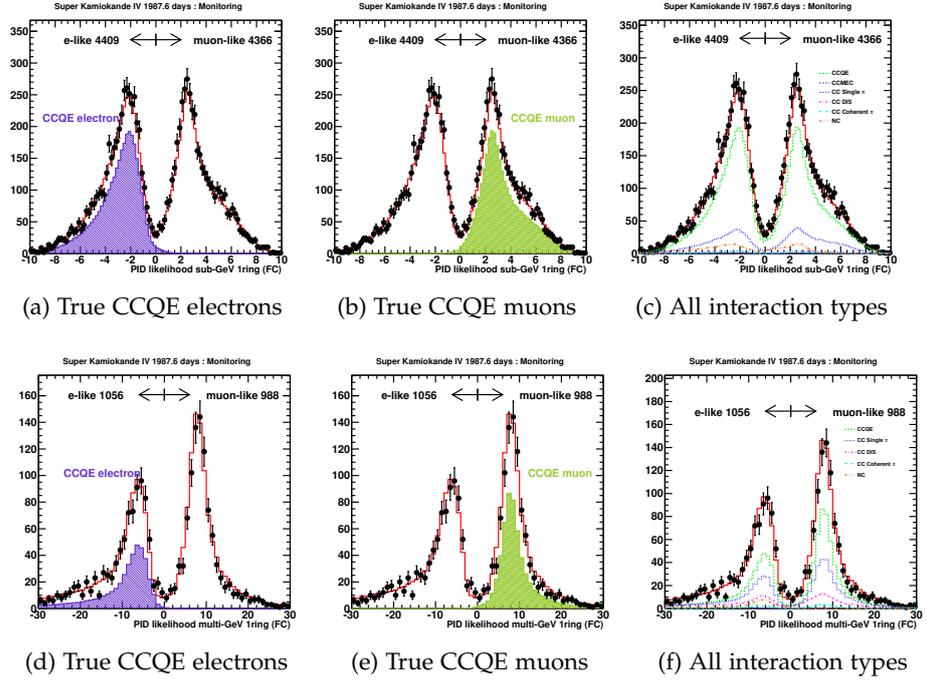
Figure 28: PID likelihood separations for the sub-GeV (top) and multi-GeV (bottom) single-ring sample for SK-IV, showing the MC (red line), and the data (points with statistical error bars). Three plots are shown overlaid with various aspects of the MC truth information.

(a) 405 MeV   (b) 1.1 GeV   (c) 4.2 GeV
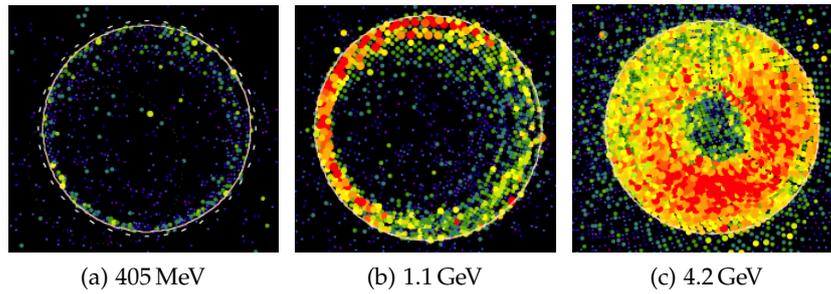
Figure 29: An example comparison between ring patterns for a muon of varying energy, where the size and color of the dots represent the hit charge of each PMT. [66]
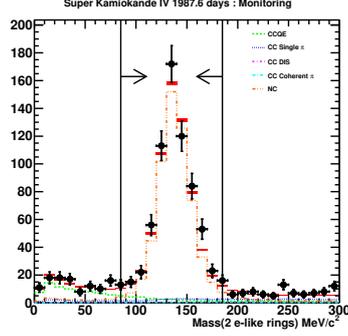
Figure 30: The π_0 mass peak after reconstruction, from the π_0-like (two fitted electron ring) sample. Both the MC with interaction type breakdown, and data points with statistical error are shown.

the ring width and Cherenkov angle gives a better measurement of the true momentum (as illustrated in Fig. 29). For single-ring events, by examination of the MC the resolution of the muon momentum on an event-by-event basis is found to be $\sim 4\%$, and for electrons is found to be energy-dependent at $\sim 10\%$, $\sim 5\%$, and $\sim 3\%$ for a 100 MeV, 500 MeV, and 2 GeV electron respectively.[11] In the case of multi-ring events, due to ring overlap the momentum resolution for each ring is decreased by roughly a factor two.

An important variable used to classify events is the "visible energy" E_{vis}, which is defined assuming that all fitted rings are *electrons* and summing their estimated momentum. This variable is used due to the simple correlation between the electron energy and the total n_{pe}, and it gives a general energy scale in units of eV which can be assigned to each event. Events are classified into so-called "sub-GeV" or "multi-GeV" samples, which are actually divided at the threshold of exactly 1.33 GeV in E_{vis}.[12]

We also define the variable E_{rec}, which conversely does take into account the fitted PID of each ring, and adds the reconstructed momenta and masses appropriately.

4.2.2.5 π_0-like event identification

Neutral pions are a common product of neutrino interactions in water (as described in Sec. 4.5.1). Events whereby only pions are detected cannot discriminate the neutrino flavour (although they do carry useful information on the neutrino's other properties), and thus are categorized differently. Since we are interested in samples with accurate

---

11 These errors are in addition to the overall energy scale error, as discussed in Sec. 3.4.
12 This is partially for historical reasons relating to the Kamiokande experiment, although is also the energy around which the leptons produced in CCQE interactions start to have good directional correlation with the incoming neutrino (which is a somewhat arbitrary statement, but one can compare e.g. the $\sim 15\%$ resolution at 1330 MeV to the $\sim 30\%$ and $\sim 60\%$ resolutions at 1000 MeV and 700 MeV respectively).

neutrino flavour, we do not plan to use such events, but by identifying them they can be removed and the flavour purity of other samples may be increased.

Fortunately, neutral pions can in principle be easily identified by their dominant decay $\pi^0 \to \gamma + \gamma$, whereby the photons produce two electromagnetic showers and thus two e-like rings in the detector, the invariant mass of which is the pion mass. Fig. 30 shows the reconstructed invariant mass of such events with two reconstructed rings; these are classified into a "π_0-like" sub-sample.

The reconstruction of two e-like rings is however not always easy; one or both rings may have low charge and be difficult to reconstruct, or the two rings may be overlapping due to a Lorentz boost between the lab frame and the decay frame. Therefore, for sub-GeV single-ring e-like events, an extra π^0 finder algorithm [92] is also applied, which performs a forced search for a second γ ring that was originally missed. The algorithm identifies events using a likelihood method, comparing the single-ring best-fit result to the two-ring best-fit, while considering also if the reconstructed invariant mass is close to the π_0 mass. These events are classified into a "single-ring π_0-like" sub-sample.

4.2.2.6 *Final division into sub-samples*

For all three of the analyses in this thesis, a strong requirement is the selection of separate ν_e and ν_μ induced samples with high flavour purity. We give here the final definitions of the FC sub-samples that will be used in all of the analyses (with the exception of Sec. 5.6, which uses instead the neutrino / antineutrino separated FC sub-sample definitions described in Sec. 4.2.3). Firstly, all π_0-like sub-samples are discarded, due to their high NC and wrong-flavor CC backgrounds, as seen in Fig. 31. All single-ring events with $E_{rec} > 100\,\text{MeV}$ for e-like events, and $E_{rec} > 200\,\text{MeV}$ for μ-like events, are selected, and multi-ring events are selected as described in the previous section.

The remaining FC events are thus generally divided by three variables (also as described in the previous section): sub-GeV or multi-GeV, e-like or μ-like, and single-ring or multi-ring. Multi-ring events are mostly multi-GeV and are not further divided by their visible energy, leaving six final FC event categories, which are:

- FC e-like single-ring sub-GeV,

- FC e-like single-ring multi-GeV,

- FC e-like multi-ring,

- FC μ-like single-ring sub-GeV,

- FC μ-like single-ring multi-GeV,

- FC μ-like multi-ring.
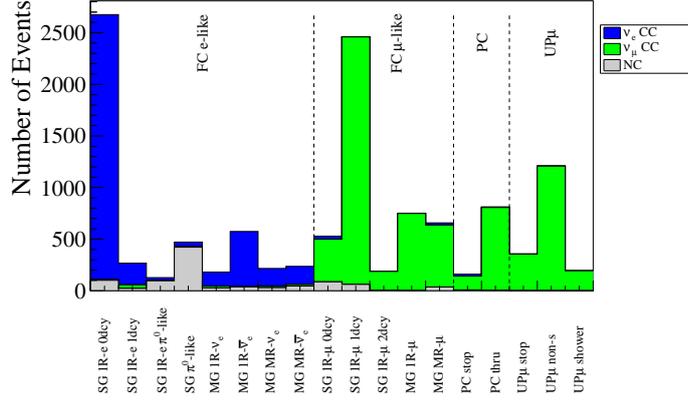
Figure 31: The flavour purity of various SK event samples, described in Sec. 4.2→Sec. 4.4, estimated by MC. Here SG and MG stand for sub-GeV and multi-GeV, while 1R and MR stand for single-ring and multi-ring, and *n*dcy refers to the number of Michel electrons detected.

4.2.3 *Neutrino / antineutrino separated selections*

As previously mentioned, the detector does not have an ability to distinguish ν from $\bar{\nu}$ on an event-by-event basis, however recent work in this area has defined and improved several statistical methods of $\nu/\bar{\nu}$ separation based on the different kinematics of their interactions. In this section, we discuss the creation of separate particle and antiparticle sub-samples for the FC dataset; since the background of wrong-sign neutrinos remains relatively high, we refer to these samples as ν or $\bar{\nu}$ *enriched* samples. We abandon for this section the abridged notation where e.g. $\nu_e \equiv \nu_e + \bar{\nu}_e$, and discuss ν_e, $\bar{\nu}_e$, ν_μ, and $\bar{\nu}_\mu$ separately.

The difference in the number of neutrino and antineutrino interactions in the detector obviously depends on two factors: the relative normalization of the two fluxes, and their interaction probabilities, as a function of energy. The true $\nu/\bar{\nu}$ ratios estimated by the HKKM group, with very small systematic errors of $O(1\%)$, are shown in Fig. 32. We make use of these theoretical ratios when defining our neutrino and antineutrino samples. Since ν are more abundant than $\bar{\nu}$, it is generally harder to make $\bar{\nu}$ samples with high purities. Many of the discussions below refer to the various neutrino interaction types, making use of the discussion in Sec. 4.5.1.

4.2.3.1 *Neutron tagging*

Free neutrons in water can be captured by hydrogen nuclei to create deuterium, with a half-life of 205 μs, by the interaction $n + p \rightarrow d + \gamma$,
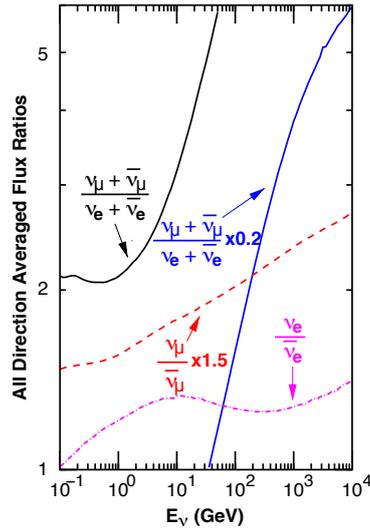
Figure 32: The $\nu/\bar{\nu}$ ratios at the SK site, as estimated in the HKKM15 simulation. [68]

where the emitted γ-ray has an energy of precisely 2.2 MeV.[13] While the detection of a low energy γ-ray in a large time window is not an easy task for a detector such as SK, as mentioned in Sec. 3.3.2 the SK-IV electronics allowed for the introduction of an AFT trigger, which records data up to 535 μs after a high-energy event. Since the primary vertex is known from the high-energy interaction, and the expected energy of the γ-ray is precisely known, it is nonetheless possible to pick out some candidates from the expected radioactive background.

An initial selection is done by first selecting candidates whereby the number of ID PMT hits n within a sliding 10 ns time window is $7 < n < 50$, where the time distribution uses the time-of-flight subtracted timing based on the APFit best-fit vertex. It is also required that the total number of ID hits within ± 10 ns of the candidate is less than 200, in order to avoid cosmic-ray muon background. Over the 535 μs period, this cut has a 33.2% efficiency to detect true neutrons, whereas an average of 4.5 false events will be detected (estimated based on dummy trigger data). These initial candidates are then fed through a neural network trained to pick out true neutron events, using 16 parameters which generally describe the spatial and timing distribution of the selected hit PMTs (mean opening angle, RMS of the time distribution, etc.). The output of the neural network has, averaging over all positions in the FV, an efficiency of 20.5% to detect true neutrons for a much-reduced background of 0.018 events per AFT trigger. More details of the algorithm can be found in [83, 84].

---

13 Neutron capture by oxygen is also possible, although capture by hydrogen is approximately 10^3 times more common.

Low-energy CCQE interactions, considering charge conservation, must proceed by one of the following:

$$\nu + n \rightarrow l^- + p$$
$$\bar{\nu} + p \rightarrow l^+ + n$$

where the left-hand nucleons are bound within a nucleus, but the right-hand nucleons are often freed. Here, the benefit of neutron tagging in neutrino / antineutrino separation is clear. At high energies however CCQE interactions are less common and the discrimination power of neutron tagging is lower.[14]

4.2.3.2 *Decay electron tagging*

As previously mentioned, it is also possible in SK to tag Michel electrons[15] originating from the decay of $\mu^\pm$. In fact, as Michel electrons have an average energy of a few tens of MeV, this is clearly much easier than tagging a 2.2 MeV γ-ray, and is reliably achieved with the standard FC reconstruction tools, although with some inefficiency for very early Michel electrons particularly for the SK-I to SK-III electronics.

To understand the benefits of electron tagging in $\nu/\bar{\nu}$ separation, an important fact to note is that while μ^- can be captured in water to form muonic atoms (approximately 20% of μ^- in SK are captured on oxygen nuclei), μ^+ has no such capture, and thus has a longer lifetime and a higher possibility to create a Michel positron. Thus, considering μ-like single-ring events (which are mostly CCQE events, especially at low energies) those with a single decay electron tagged should contain more μ^+ and thus a higher fraction of true $\bar{\nu}_\mu$.

However, considering medium-energy neutrino interactions that create meson showers, we have the following possible interactions:

$$\nu + p \rightarrow l^- + N + P^{++}$$
$$\nu + n \rightarrow l^- + N + P^+$$
$$\bar{\nu} + p \rightarrow l^+ + N + P^0$$
$$\bar{\nu} + n \rightarrow l^+ + N + P^-$$

where N is the modified outgoing nucleus, and P represents a meson shower (generally mostly pions) with the superscript indicating the total charge. Considering the decay chain $\pi^\pm \rightarrow \mu^\pm \rightarrow e^\pm$, as given in Eq. (29) and Eq. (29) on page 19, it is seen that for neutrino events causing hadronic showers, ν_e and ν_μ events would tend to have

---

14 As a final note on the topic of neutron tagging, studies evaluating the possibility to add small amounts of gadolinium in SK and future detectors [93, 94] are recently nearing completion; this greatly increases the capture and identification rate of neutrons, and is expected to be a major upgrade for neutrino water Cherenkov detectors.

15 And of course positrons, however with no power to separate between the two.
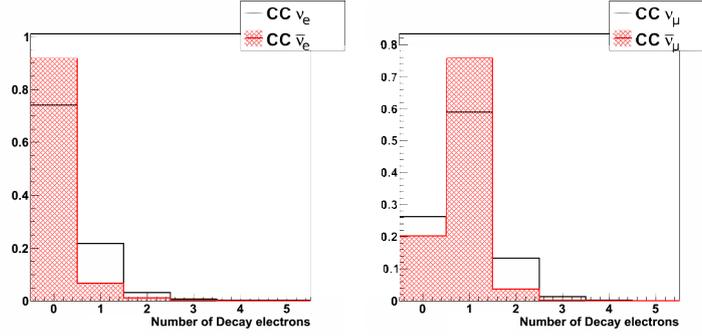
Figure 33: The true neutrino purities of the multi-GeV single-ring e-like (left) and μ-like (right) samples, as a function of the number of tagged Michel electrons (or positrons). The histograms are normalized to unity, and the wrong-flavour background of each sample is not plotted. [84]

more decay electrons. This effect should be expected in the multi-ring sample, however it can leak into the single-ring sample if the meson shower is below the Cherenkov energy threshold; in this case, it competes with the opposite effect described in the previous paragraph (where $\bar{\nu}_\mu$ events are more likely to have a decay electron). The prediction from MC is shown in Fig. 33, where we see the result of the two effects for multi-GeV single-ring events.

4.2.3.3   *Hadronic multiplicities*

Experimental data indicate [95] that ν interactions generally have a higher multiplicity than $\bar{\nu}$ interactions, which is to say that they tend to create more particles in hadronic showers. Thus, events with a higher number of fitted rings have a slightly higher chance to be a true ν.

There are also secondary effects of the increased multiplicity for ν interactions. As the deposited energy is shared amongst more rings, the fraction of E_{rec} in the primary ring to the total E_{rec} tends to be slightly smaller. The distribution of decay electrons within the tank is also affected; as more particles are involved in the shower, the chances of a decay electron to be observed further from the primary vertex (both spatially and temporally) is increased. Finally, the PID variable is often shifted slightly towards e-like instead of μ-like, originating in the fact that there is a higher chance for unidentified mesons to interfere with the primary ring identification (i.e., to increase the "fuzziness" of the primary ring).

4.2.3.4   *Final division into sub-samples*

As in the previous section, we divide events as single-ring or multiring, and for the single-ring samples also as sub-GeV or multi-GeV. The π^0 samples are also separated in the same way as before. All

sub-GeV single-ring events with $E_{rec} > 100\,\text{MeV}$ for *e*-like events, and $E_{rec} > 200\,\text{MeV}$ for *μ*-like events, are selected – however unlike the previous section, they are also divided by the number of decay electrons n_d that were detected. For single-ring multi-GeV and multi-ring events, a more complicated procedure is used in to divide both groups into four ν_e, $\bar{\nu}_e$, ν_μ, and $\bar{\nu}_\mu$-like samples, as described below.

Single-ring multi-GeV e-like events are classified as ν_e-like if $n_d > 0$ and $n_n = 0$ (where n_n is the number of tagged neutrons, and is only available and checked in SK-IV), otherwise they are classified as $\bar{\nu}_e$-like.

Single-ring multi-GeV μ-like events are classified as ν_μ-like or $\bar{\nu}_\mu$-like based on an artificial-neural-network likelihood method,[16] which considers the parameters n_d, n_n, the PID likelihood, the maximum Michel electron distance from the primary vertex, and the longest time between the primary event and a Michel electron. This neural network, and others mentioned in this section, were defined in [89] and further developed in [84]; these two PhD theses contain detailed information such as the output shape of the network, the choices of final cut parameters, and so on.

Multi-ring e-like events are classified as ν_e-like or $\bar{\nu}_e$-like based on a neural network considering n_d, n_n, the PID likelihood, the number of rings, and the primary ring's E_{rec} compared with the total E_{rec}.

Multi-ring μ-like events are classified as ν_μ-like or $\bar{\nu}_\mu$-like based on a neural network considering all of the parameters mentioned above.

Finally, any multi-ring multi-GeV event that fails the likelihood selections is classified into a "multi-ring other" sample. From the above definitions, using the SK MC (discussed in Sec. 4.5) we show the estimated purities of each of these sub-samples in Tab. 6. We see that each sample has a distinctive shape in the distribution of its true neutrino content. For ν_e and ν_μ, it is possible to obtain several samples with purity $> 70\%$, whereas for $\bar{\nu}_\mu$ we can manage at best 50%, and for $\bar{\nu}_e$ only around 30%. In Sec. 5.6, when these samples are used for analysis, we will discuss a sub-selection and binning of these events.

---

16 An artificial neural-network is a multivariate analysis method, which can be trained to discriminate between a set of possible outputs given certain input parameters. The end result is somewhat similar to a likelihood method, except that the weights contributed by the individual parameters are automatically tuned by the neural network, and can depend on the values of the other parameters.

Sub-sample	Estimated fractional purity					Events
	ν_e	$\bar{\nu}_e$	ν_μ	$\bar{\nu}_\mu$	NC $+ \nu_\tau + \bar{\nu}_\tau$	
1. single-ring sub-GeV e-like $n_d = 0$	**73.4**	23.2	0.3	0.1	3.1	3107
2. single-ring sub-GeV e-like $n_d \geq 1$	**72.4**	2.1	13.0	3.9	8.5	344
3. single-ring π^0-like	12.9	4.3	1.5	0.2	**81.0**	183
4. single-ring sub-GeV μ-like $n_d = 0$	3.2	1.0	**74.4**	8.6	12.8	829
5. single-ring sub-GeV μ-like $n_d = 1$	0.1	0.0	**68.9**	29.2	1.8	4114
6. single-ring sub-GeV μ-like $n_d \geq 2$	0.0	0.0	**94.0**	4.0	2.0	314
7. two-ring π^0-like	6.9	2.3	1.6	0.1	**89.2**	573
8. single-ring multi-GeV ν_e-like	**74.7**	7.9	9.7	2.6	5.1	102
9. single-ring multi-GeV $\bar{\nu}_e$-like	**54.8**	32.9	2.7	0.6	9.0	833
10. single-ring multi-GeV ν_μ-like	0.3	0.2	**72.9**	26.4	0.2	670
11. single-ring multi-GeV $\bar{\nu}_\mu$-like	0.0	0.0	**50.3**	49.6	0.1	600
12. multi-ring ν_e-like	**54.9**	12.2	11.5	0.9	20.6	328
13. multi-ring $\bar{\nu}_e$-like	**54.2**	29.5	4.3	0.6	11.4	235
14. multi-ring ν_μ-like	3.1	0.5	**77.8**	12.4	6.2	645
15. multi-ring $\bar{\nu}_\mu$-like	0.7	0.0	**63.5**	34.7	1.0	447
16. multi-ring other	24.2	2.7	**39.0**	4.8	29.3	575

Table 6: The fractional purities of each sub-sample defined for the $\nu/\bar{\nu}$ separated analysis and expected number of events, estimated by MC, where n_d is the number of decay electrons. In this table, samples that will later be discarded are also included, in order to examine the relative purities of each sample. This table refers to SK-IV only, however selections in other periods are similar within 2 ~ 3%. The bold text shows the primary component of each sample.

4.3 PARTIALLY CONTAINED SAMPLE

4.3.1 *Reduction*[17]

The PC reduction is (like the FC sample) divided into five steps, and is denoted PC1→PC5. Considering that PC events by definition have OD activity, which is expected to be generated by a single exiting muon track, the reduction process must be fundamentally quite different, even at the early steps. Cuts are described in the same way as in Sec. 4.2.1.

4.3.1.1 *PC1*

The first PC reduction consists of some basic cuts:

1. total ID $n_{pe} > 1,000$

which is roughly equivalent to the energy deposited by a muon of initial momentum $310\,\text{MeV}$ traveling $\sim 2\,\text{m}$, which is used as it is the minimum energy for a single muon created in the FV to penetrate into the OD. Next,

2. width of the OD timing distribution $< 260\,\text{ns}$

which references the RMS of the OD hit timings, and should remove events with two event clusters (such as a through-going muon that enters and later exits the detector). This cut is supported by a simple positional cluster detection algorithm searching for nearby hits, requiring

3. number of hit clusters (simple algorithm) in the OD = 1.

4.3.1.2 *PC2*

The second reduction depends on a more advanced clustering algorithm, which divides the ID PMTs into a grid of 441 patches, and the OD into 121 patches, and searches for clusters by finding steep gradients between neighbouring patches. The similar cut

4. number of hit clusters in the OD ≤ 1

is performed, and if an OD cluster is found it is also required that

5. total ID n_{pe} in the ID cluster closest to the OD cluster (within $200\,\text{cm}$ of the highest-charge PMT) is $> 1,000$.

---

17 This section describes the reduction process as used in SK-I, in order to show a broad understanding of the PC data processes. Several tunings and updates were later applied, particularly when OD segmentation was installed at the beginning of SK-III, optically separating the top, center, and bottom areas of the OD. For a thorough description of the SK-IV process, see [89], and for more detail on SK-I and SK-II please see [96].

Next, it is required that

6. least active detector region has < 7 hit PMTs

where the "detector region" refers to either the top, side, or bottom surfaces of the cylindrical SK detector. These cuts are again designed to find PC events that have a single exit point in the detector, and are related to an ID hit cluster.

4.3.1.3   PC3

The third reduction targets the specific backgrounds of flasher events and stopping muons. The long time distributions of flasher events are cut by requiring:

7. minimum number of OD hits in a sliding 100 ns time window from +0 ns to +500 ns is < 10 (this criteria is relaxed to < 15, if more than 800 ID PMTs are hit).

Stopping muons are then rejected using the point-fit algorithm, which roughly gives the ID vertex and direction. By extrapolating backwards, a rough entry-point can be found in the OD, which should contain an early hit cluster if the event was a stopping muon:

8. number of hit OD PMTs within 8 m of the estimated entry-point < 11.

4.3.1.4   PC4

Some stricter cuts on through-going cosmic-ray-muons, which by chance had low OD activity at the entry point, are still necessary in the fourth reduction. This is done using a specialized through-going muon fitter, which uses the most likely early hit ID cluster. Events with

9. goodness-of-fit of through-going muon fitter < 0.85

are accepted, whereby the entry cluster's fitted position and timing will correspond poorly to the projected through-going muon track. A similar-purpose cut is made using the point-fit information, comparing the fitted direction to the earliest hit ID PMT, requiring

10. inner product of point-fit vector and earliest-hit vector > -0.85,

i.e. these two vectors should not be back-to-back.

4.3.1.5   PC5

After the fourth reduction, the event rate is down to ~ 20 events per day, and the application of advanced fitters[18] to these events is feasible. The fifth reduction then uses several finely-tuned criteria based

---

18 Including `apfit`, `stmfit`, `pfdofit`, and `msfit`.

	EVENT RATE $[\text{DAY}^{-1}]$		EFFICIENCY [%]	BACKGROUND [%]
	DATA	MC		
SK-I	0.67 ± 0.02	0.69	85 ± 2	–
SK-II	0.63 ± 0.03	0.67	80 ± 5	–
SK-III	0.66 ± 0.04	0.69	96 ± 1	2
SK-IV	0.66 ± 0.02	0.69	95 ± 1	2

Table 7: The final daily event rates of events passing the PC reduction, with statistical error (negligible for MC). The meaning of the "efficiency" and "background" columns is the same as in Tab. 5, as was explained at the end of Sec. 4.2.1. In the PC case however, eye-scanning records (for estimating the background rate) are only available for SK-III and SK-IV. Values are taken from [76].

on these reconstructed tracks to target the remaining background sources. This includes many types of through-going and stopping muon cuts; targeting for example muons that pass through the detector superstructure and PMTs, or corner clipping muons. The fifth reduction also targets veto-counter events and electronics noise events, in a similar way to the FC reduction.

4.3.1.6 *Final reduction*

The final requirements to pass the PC reduction, in SK-I, are:

11. the distance from the neutrino interaction best-fit vertex and the wall is $> 2\,\text{m}$,

12. the number of OD PMTs found by a clustering algorithm is ≥ 16,

13. the visible energy $E_{vis} > 350\,\text{MeV}$.

The final PC sample has a somewhat lower efficiency for true neutrino events than the FC sample, as shown in Tab. 7. Of course, this should not bias the analysis, since the MC events are selected with the same efficiency.

4.3.2 *Reconstruction*

The PC sample uses the same software and essentially the same techniques as the FC case, with some modifications to account for the extra charge deposited in the OD. Most importantly, events are divided into two samples:

- PC stop,

- PC through,

based on whether the muon particle stops inside of the OD or leaves the detector completely (where the PC through sample has much larger statistics). This separation is done because the momentum of the muon can be determined accurately for events in the PC stop sample, but for the PC through sample the energy deposited by the particle after it has left the detector cannot be measured, and the momentum resolution is much worse.

The selection of PC *stop* events is made in a simple fashion [97], by testing the observed charge in the OD, requiring

1. maximum OD n_{pe} in a sliding $500\,\mathrm{ns}$ time window $< \frac{2}{3}n_{\mathrm{exp}}$,

2. at least one of the two highest energy rings in the ID is classified as μ-like,

where n_{exp} is the expected deposited n_{pe} of a muon track, assuming it exited the OD (which varies as a function of the fitted exit point, to take into consideration the geometry of the detector). All other events are categorized into the PC *through* sample.

4.4   UPWARD-GOING MUON SAMPLE[19]

4.4.1   *Reduction and reconstruction*

The UPMU reduction has the task of distinguishing upwards-going muons that enter the detector from below, originating from a neutrino interaction in the rock, and either stop inside the detector or exit through the top or sides. Like the FC and PC reductions, background data are removed by a series of automated cuts, however unlike the FC and PC reductions, the UPMU reduction also uses subtractions of the estimated backgrounds, and manual eye-scanning to select final events. The reduction consists of four automated steps (UM1→UM4) and a final eye-scanning step.

4.4.1.1   *UM1 to UM4*

The first reduction keeps only events with a total ID $n_{\mathrm{pe}} > 6000$ (for SK-II, > 3000), which corresponds roughly to a $1\,\mathrm{GeV}$ muon traveling a few meters. It also saves by default all ultra-high energy events with total ID $n_{\mathrm{pe}} > 1.75 \times 10^6$, which occur only once every few days, and are used for astrophysical studies [101, 102, 103, 104]. These "astrophysical" events require a completely different reconstruction process, and are not used in this analysis.

Events are then passed to a series of fitters of increasing complexity. UM1 uses the fitter `stopmulst`, which does some simple cluster searching, and `muboy`, which provides a direction and goodness-of-fit

---

19 more details on the upmu sample can be found in the theses [98, 99], or the paper [100].

for the muon track. Events are passed through with some fairly relaxed criteria on the goodness-of-fit and direction; allowing up-going and near-horizontal single-track events, but cutting events that are clearly down-going or multi-track events. Events that pass close by the seams between the top or bottom cap and the side of the detector ("corner clipping" events) are also cut, as they are difficult to reconstruct and can cause false UPMU events (corner clipping events should in principle not make the final sample, due to a requirement that the track length be at least $\sim 7\,\text{m}$, described later).

Next, UM2 applies a series of more complicated fitters, classifying events as either "stopping" (where the muon comes to rest inside the detector) or "through-going" (where the muon exits the detector). Some tighter cut criteria target background events using the fitter information. UM3 applies the most precise and processing-intensive fitter, called `precisefit`. For SK-IV only, an additional UM4 step was added, which make use of the additional OD segmentation installed at the beginning of SK-III (optically separating the top, bottom, and side regions). OD events that cross this region are examined carefully, to further reduce the number of corner clipping events.

An event need not be reconstructed perfectly as an UPMU by all fitters, but their outputs are considered together to reach a decision, and establish a final entry point, momentum, directional vector, and track length. Events that pass the selection are also flagged into two sub-samples, "stopping" or "through-going", depending on the fitter information..

4.4.1.2 *Eye scanning*

After UM4, the UPMU data still contains $\sim 20\%$ background events. In many of these cases, the fitting software has landed in a local minima or otherwise produced a poor fit; for example due to events near the detector walls or corners (causing difficult timing distributions), or large regions of PMTs reaching their charge saturation point, or mis-selection of PMT hit clusters. A visual examination of the ID and OD charge patterns and the fitted Cherenkov ring shapes can often quickly reject such badly fitted events,[20] and every UPMU event is checked by two expert scanners who tag background events; disagreements between scanners are rare, and the tagging of true UPMUs as background is estimated to be negligible or even zero.

4.4.1.3 *Final reduction*

Events passing the eye-scanning stage have a single fitted muon track, and are flagged as stopping or through-going. The final selection criteria for the stopping sample requires

---

20 Thanks in general to the evolutionary process, and the superior pattern-matching ability of the human.

(a) Stopping muons

(b) Through-going muons
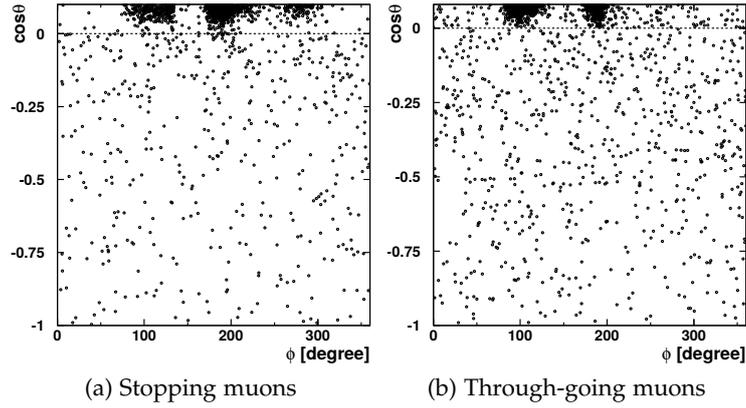
Figure 34: Fitted stopping and through-going muon data for SK-I, where ϕ is the azimuthal angle and θ is the zenith angle. Mostly up-going ($\cos\theta < 0$) neutrino-induced events are shown, although in the $0 < \cos\theta < 0.1$ range some clusters of the cosmic-ray background are clearly visible.

1. fitted muon momentum $> 1.6\,\text{GeV}$,

2. number of hit OD PMTs within $8\,\text{m}$ of the best-fit exit point is < 10 (for SK-II, < 16),

where the first cut eliminates short tracks (less than $\sim 7\,\text{m}$) which are background rich, and the second cut introduces a tight requirement that only selects stopping events. The through-going sample events are similarly required to have

1. distance between the best-fit entry and exit points of the ID $> 7\,\text{m}$,

2. number of hit OD PMTs within $8\,\text{m}$ of the exit point is ≥ 10 (for SK-II, ≥ 16).

The final efficiency of the UPMU reduction is calculated based on MC, considering only true up-going events with a vertex outside the ID and a path length of $> 7\,\text{m}$, and is shown in Tab. 8.

4.4.1.4 *Background estimations*

The UPMU sample has a non-removable background, due to cosmic-ray muons coming from near horizontal directions that scatter before entering the detector and become up-going, which cannot be distinguished from neutrino-induced up-going muons. These deflections are rare, but the backgrounds are non-negligible considering that the cosmic-ray muon flux (even at the SK detector $\sim 1\,\text{km}$ underground) is higher than the true neutrino UPMU flux by a factor of $\sim 10^5$.

Due to the shape of the mountain, the rock overburden from the detector varies strongly depending on the azimuth and zenith angles;

	EVENT RATE $[\text{DAY}^{-1}]$		EFFICIENCY [%]	BACKGROUND [%]
	DATA	MC		
SK-I	1.41 ± 0.03	1.25	99.0 ± 1.0	2.7 ± 0.5
SK-II	1.33 ± 0.04	1.23	97.8 ± 1.0	4.6 ± 0.8
SK-III	1.49 ± 0.05	1.24	99.1 ± 1.0	3.0 ± 0.6
SK-IV	1.38 ± 0.09	1.23	99.0 ± 1.0	2.4 ± 0.3

Table 8: The final daily event rates of events passing the UPMU reduction, with statistical error (negligible for MC). The meaning of the "efficiency" column is the same as in Tab. 5. The meaning of the "background" column is described in the text of this section. The values are taken from [89] for SK-I to SK-III, and [76] for SK-IV.

in Fig. 34, some cosmic-ray muon clusters can be clearly seen just above the horizontal, at certain azimuthal angles. This azimuthal distribution (in the $0 < \cos\theta < 0.1$ range) is extrapolated downwards in zenith by an exponential function, and fitted against the expected neutrino-induced muon distributions. The final data set is then given as a distribution, defined as the observed data minus the fitted backgrounds, as a function of zenith angle and reconstructed momentum.

The final UPMU sample efficiencies are shown in Tab. 8. The data rates are given before background subtraction, and the "background" column represents the subtracted estimated background, and the systematic errors on the background subtraction process. These errors are generally $10 \sim 20\%$ of the subtracted background amount, and come primarily from the statistical uncertainties in the background fitting procedure.

4.4.2 *Sub-sample categorization*

So far, UPMU events were selected, reconstructed, and divided into stopping and through-going samples. The through-going events are further divided into two sub-samples depending on whether or not they undergo intense electromagnetic showering, which occurs when the muon's energy approaches $1\,\text{TeV}$, due to increased pair production and bremsstrahlung emission relative to simple ionization at these energies.

The showering sample is thus selected by requiring a greater energy deposited per unit range along the fitted muon track. The parameter used is a χ^2-inspired statistic, based around the observed charge Q_i associated with each $50\,\text{cm}$ ID track segment i, defined as

$$\chi^2 = \sum_{i=3}^{N-2} \left[\left(\frac{Q_i - Q_{\text{ave}}}{\sigma_{Q_{\text{ave}}}} \right)^2 + \left(\frac{Q_i - Q_{\text{exp}}}{\sigma_{Q_{\text{exp}}}} \right)^2 \right] \qquad (47)$$

where $Q_{\text{ave}} \equiv \langle Q_i \rangle$, and Q_{exp} is a constant representing the expected Q_{ave} for a non-showering muon. Furthermore, $\sigma_{Q_{\text{ave}}}$ and $\sigma_{Q_{\text{exp}}}$ are the estimated errors on those parameters.[21] This parameter thus prefers tracks with the sudden bursts or generally high values of energy loss associated with a showering muon track.[22] The sum ignores the first and last two track sections, which are within 1 m of the entry and exit points, where a small error in the reconstructed track could cause a large change in the associated values of Q_i.

The final sub-sample definitions of the UPMU events are thus:

- UPMU stop,

- UPMU showering,

- UPMU non-showering.

Stopping events are clearly lower energy (considering that they do not have enough energy to fully traverse the detector) – thus the three samples provide some energy discrimination for the UPMU events, as was shown in Fig. 18. Of course, the amount of energy deposited by the muon before entering the detector is completely unknown, so while we can reconstruct a muon momentum at the point of entry, an estimation of the true energy spectrum of UPMU neutrinos can be done only in a statistical sense by comparison to MC.

4.5   MONTE CARLO

In the SK Monte Carlo (MC), neutrino interaction events are generated from an interpolation of the energy and azimuthal distributions of the HKKM07 [18] atmospheric flux model, and the NEUT [105] neutrino interaction simulator. NEUT calculates the results of the primary neutrino interaction with a water molecule and all secondary nuclear effects, and then the list of exiting particles is handed off to the detector simulation SKDetsim, which handles all propagation and physical effects necessary in water Cherenkov detectors, and fully simulates the detection and digitization processes of the PMTs and SK electronics. After simulation of the events, reduction and event reconstruction processes are applied, which are essentially identical to those applied on the data. The MC is generated in all four detector configurations, corresponding to 500 years of livetime for each of the SK-I to SK-IV periods.

---

21 For example, we expect $\mathrm{d}E/\mathrm{d}x = 2\,\text{MeV}\,\text{cm}^{-1}$ for a 1 GeV muon; considering the average expected non-showering muon energy, we can estimate the average $\mathrm{d}E/\mathrm{d}x$ and its spread, and thus Q_{exp} and $\sigma_{Q_{\text{ave}}}$. The error $\sigma_{Q_{\text{ave}}}$, on the other hand, is a statistical one based on the Q_i.

22 The actual selection uses an additional parameter accounting for detector tunings, and fitted to a MC prediction (see [98]).
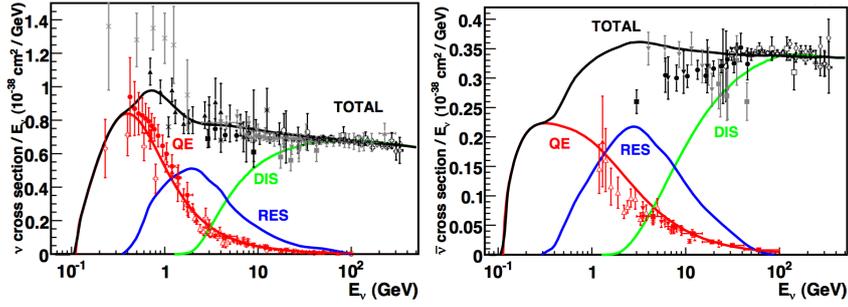
Figure 35: Cross-sections for the different types of neutrino interaction, showing ν_μ (left) and $\bar{\nu}_\mu$ (right). The solid lines are predictions are from the NUANCE neutrino interaction generator, and the data points are taken from various experiments. The meanings of QE, DIS, and RES are described in the text of this section. [106]

INTERACTION TYPE	FINAL PRODUCTS
(quasi-)elastic scattering	$\nu + N \rightarrow l + N'$
single π production	$\nu + N \rightarrow l + N' + \pi$
single γ production	$\nu + N \rightarrow l + N' + \gamma$
single K production	$\nu + N \rightarrow l + \Lambda + K$
single η production	$\nu + N \rightarrow l + N' + \eta$
coherent π production	$\nu + O \rightarrow l + X + \pi$
deep inelastic scattering	$\nu + N \rightarrow l + N' + hadrons$

Table 9: Primary neutrino interactions considered by NEUT. Here, N and N' are the initial and final nucleons (e.g. n or p), X is the final nucleus, and other notations are standard. Each interaction can proceed by both charged or neutral currents, which is to say that the outgoing lepton l can be either a neutrino or a charged lepton ($e^\pm$, $\mu^\pm$, $\tau^\pm$), if the appropriate choice is also made for N'.

4.5.1 *Neutrino interaction simulation*

4.5.1.1 *Primary interactions*

The neutrino interaction generator NEUT considers several types of primary neutrino interaction, as listed in Tab. 9, with cross-sections estimated based on theoretical models with free parameters fitted by a wide array of experimental data, such as bubble chambers and other neutrino experiments. Several target materials such as carbon, argon, iron, oxygen, and hydrogen are implemented in NEUT (the latter two of which are of interest in SK). The cross-section of a neutrino interaction with an electron is several orders of magnitude smaller than those with nuclei, and is neglected.[23]

---

23 An excellent overview of neutrino interactions is given in [107].

QUASI-ELASTIC SCATTERING (QE)   is the dominant process up to around $1\,\mathrm{GeV}$, whereby a neutrino scatters off an entire nucleon, usually ejecting one or even several nucleons from the nucleus. This is often referred to as "elastic scattering" in the NC case, and "quasi-elastic" in the CC case. The cross-sections in NEUT are based on the Llewellyn-Smith model [108], for free nucleon scattering. To obtain the QE cross-sections for particular nuclei, modifications by the the Smith and Moniz [109] model are used, where a neutrino is taken to interact with a bound nucleon described within the nucleus by a relativistic Fermi gas model (although simple free nucleon interactions can still occur, at a ratio predicted by the model). These models contain both vector and axial-vector form factors (which are necessary due to the $V - A$ structure of the weak interaction), both of which are modeled as "dipole" type,[24] which is to say the form factors F depend on the total four-momentum transfer q^2 as

$$F(q^2) = F_0 \left(1 - q^2/M^2\right)^{-2} \tag{48}$$

where F_0 is a base value predicted by the model, and M is the mass parameter, which is not predicted but must be determined by experimental measurements. For the vector form factor, the mass parameter $M_{\mathrm{QE}}^{\mathrm{V}}$ is well determined by electron scattering experiments, whereas the axial-vector mass parameter $M_{\mathrm{QE}}^{\mathrm{A}}$ is set at $1.21\,\mathrm{GeV}$ based on data from the K2K [112] and MiniBooNE [113] neutrino experiments (here, the subscript QE is added in order to differentiate this parameter with other axial-vector mass parameters that are used in e.g. the single meson production interactions, which are based on different models and thus may have different values). The systematic errors on the QE scattering cross-section are estimated using the errors on the experimental results measuring the mass parameters, and the difference between available theoretical models (for example, between the Smith & Moniz model, and the Nieves [114] model of nuclear interactions). As mentioned before, the QE process is particularly useful for measuring neutrino interactions due to the simple two-body kinematics, as the recoil of the final nucleon is generally far below the Cherenkov threshold in SK.

RESONANT PRODUCTIONS (RES)   are interactions producing a single meson (or photon), whereby at specific energies excited baryonic states of nucleons are created within the nucleus, which then proceed to decay into nucleons and mesons. Of these interactions, single π production is the dominant process. In NEUT the cross-sections for

---

24 In fact, this type of form factor disagrees with experimental measurements for the vector coupling [110]. It is nonetheless maintained in NEUT, partially for reasons of comparability with other generators and older experiments. Recent measurements also indicate that the axial vector coupling may need to be updated in the future [111].

these interactions are mostly based on the Rein and Sehgal model [115], which is a calculation of all possible baryonic excited states up to an invariant mass of $W = 2\,\text{GeV}$ along with their decay mode probabilities. Similar to QE interactions, the RES model also depends on experimentally-measured mass parameters, mostly fitted according to the MiniBooNE data. So-called "pionless delta decay", where the energy of the baryonic resonance may be adiabatically absorbed by the nucleus, is also modeled based on [116] (accounting for $\sim 20\%$ of excited resonances). These resonance interactions are important around the few-GeV range.

In coherent π production, the neutrino interacts with the nucleus as a whole and produces a forward-going pion (or very occasionaly, a heavier meson) without transferring a large amount of momentum to the nucleus, such that no excited nuclear states or ejections of nucleons occur within the modified nucleus (denoted X in Tab. 9). The angular distribution of the outgoing lepton and meson are strongly peaked in the forward direction. Coherent π production is predicted to occur in both CC and NC interactions by the Rein and Sehgal model [117], and while both have been seen in many high energy experiments [107], measurements by K2K [118] and SciBooNE [119] are consistent with the absence of such CC interactions at the 90% confidence level at low energy (below $2\,\text{GeV}$), and a conservative systematic error of 100% is assigned to the predicted cross-section. For NC interactions, a systematic error of 30% is assigned based on the SciBooNE [120] measurement.

DEEP INELASTIC SCATTERING (DIS)   interactions are important at higher energies, where neutrinos can resolve the quark structure of the nuclei, and cross-sections must be calculated considering this structure [121, 122]. These interactions create complicated hadronic outputs, where the particle multiplicities for a given energy are estimated by a study of bubble chamber data [123, 124] for $W < 2\,\text{GeV}$, and simulated by the package PYTHIA/JETSET [125] above that. The fractional systematic error on the total cross-section for all DIS interactions is estimated, by comparison to the CC inclusive cross-section measurement by MINOS [13], as $0.4 \times (E_v[\text{GeV}])^{-1}$, e.g. 40% at $1\,\text{GeV}$ but decreasing to 4% at $10\,\text{GeV}$.

4.5.1.2 *Final state interactions*

It is known that the mesons and hadrons produced in a "primary" interaction generally interact within the nucleus before escaping, and NEUT also tracks these nuclear effects (known as final state interactions or FSI) until all generated particles have been reabsorbed into the nucleus, or have traveled far enough to be considered to have left the nucleus. This is done by the Monte Carlo cascade model, be-

ginning all particles at the vertex of their generating interaction,[25] and stepping each particle by a certain distance while calculating the probability to undergo various interactions at each step. The nucleon density is based on the Woods-Saxon distribution.

In general the implementation of all possible FSI effects is an extremely difficult and ongoing task, based on theoretical models but requiring many re-tunings to fit the experimental data. While NEUT is often updated to keep consistent with the current data, the uncertainties in the final outputs are relatively high, and the errors on the various cross-sections and FSI effects must be thoroughly considered as systematics in the final neutrino analyses.

4.5.2 *Detector simulation*

The detector simulation SKDetsim is based on a customised version of GEANT3 [126]. It contains a full model of the detector, and incorporates all calibrations and known effects discussed so far in this thesis – including precise measurements of the PMT and detector geometries, water properties, reflection properties of the various materials, electronics, and PMT response (such as angular acceptance, quantum efficiency, and charge-timing response).

Some models of the physical processes relevant at SK have also been enhanced or updated. All charged particles in the simulation emit Cherenkov radiation, and all standard electromagnetic effects such as ionization, pair production, Compton scattering, Bremsstrahlung, and so on, are included. Particle decays, such as the Michel decay for muons, are fully simulated. Upon further nuclear collisions, hadronic interactions are simulated by the CALOR package [127] above an energy of 500 MeV, whereas below this particles are handed back to NEUT and treated by the cascade model, the combination of which gives the best reproduction of the current SK data. Recent updates to SKDetsim (and NEUT) are described in [128].

4.5.3 *Event weights*

Each neutrino event has several associated reweighting factors. Firstly, events are reweighted from the HKKM07 atmospheric flux model to the more recent HKKM11 [19] model on an event-by-event basis. Weights due to neutrino oscillation are also calculated and applied per event; SK analyses (except the oscillation analysis itself) generally use the three-flavor oscillation parameterization taken from the yearly PDG average [59], under the assumptions of normal hierarchy ($\Delta m_{32}^2 > 0$) and CP symmetry ($\delta_{CP} = 0$). Finally, weights are applied to account for the expected neutrino flux changes due to changes in the

---

25 Except in the case of an initial DIS interaction, where hadrons may be generated away from the neutrino interaction vertex.

solar activity. The HKKM11 model provides a complete dataset for three particular degrees of low, medium and high solar activity. The average degree of solar activity for each SK period was estimated from cosmic neutron count data [129], and all events for a given SK period are assigned the average solar weighting taken from a linear combination of the three datasets.

Although the HKKM11 model is taken as the default one in most SK analyses, it is possible to reweight each event instead based on the published energy spectrum of other flux models, such as the Fluka [20] or Bartol [21] models. In general however, these groups have not released detailed directional distributions for the neutrino flux at the SK site, and those are always taken from the HKKM07 model.

4.5.4 *Systematic Error Estimation*

While not strictly a part of the MC simulation, we discuss here the method of calculation of systematic errors on the SK data, coming from uncertainties on the detector reconstruction procedures, and the cross-section and oscillation models. A systematic error database is maintained by the SK collaboration, which is a codebase from which the effects of the systematics may be calculated on an analysis-by-analysis basis. This database is the result of thorough systematic error studies over the > 20 year operational history of the detector, and is considered to give an accurate estimation of all known systematic error sources on the SK data. All analyses in this thesis use the above database, although in some cases additional systematics that were not previously considered are checked and estimated where necessary.

The actual method used by the software is as follows. Firstly, the software requires a defined output binning of reconstructed energy E_{rec} and reconstructed zenith angle θ, for each sub-sample being used; we denote all bins in this three-dimensional structure by the index j. Importantly, this binning can differ depending on the needs of a particular analysis. Each MC event is then input to the software, which carries the information of the neutrino energy E_ν, true zenith angle, and which sub-sample the event was reconstructed into by the fitting procedure. The software then calculates, for each event and each systematic error, the estimated change in each output bin j for a 1σ change in the given systematic.

As a simple example, we can consider a single-ring μ-like event with no identified Michel electron ($n_d = 0$). This event would have a few-percent chance to migrate into the $n_d = 1$ sample, when considering the "Michel electron tagging efficiency" systematic error source. It would also have few-percent chance to migrate into a different E_{rec}

	LIVETIME [DAYS]				
	SK-I	SK-II	SK-III	SK-IV	TOTAL
FC AND PC	1489	799	518	1993	4799
UPMU	1646	828	636	1993	5103

Table 10: Livetimes for each SK period. UPMU data livetimes are longer due to less strict conditions for good run selection.

bin, considering the energy scale error[26]. Thus, the value of the output bin j in which the event is classified would be decreased by a few percent, and some other bins would gain a value equivalent to the decrease, for those two systematic errors. After looping over all events, the total effect of all errors (indexed as k) in all bins j is given by the set of coefficients f_{jk}, which represent the fractional shift in each bin for the 1σ effect of each error.

A full list of all systematics included in the database is given in Appendix A, with some descriptions of the meaning of each systematic when necessary. For the detector related errors, 29 sources are considered and estimated for each SK period (a total of 116 errors), the cross-section uncertainties are accounted for by 17 errors, and finally there are 6 oscillation parameter uncertainties. The detector errors have been estimated by a combination of MC studies, and physics studies on control samples such as atmospheric muons or calibrations as described in Sec. 3.4. The neutrino cross-section errors are taken from the model uncertainties as discussed in Sec. 4.5.1, and the oscillation parameter uncertainties are based on the results from other experiments - in this thesis, we update the oscillation errors to use the latest parameterizations from the PDG [59].

4.6   DATA AND MONTE CARLO SETS

Tables 86 and 11 summarize the detector livetimes and event numbers in the data and MC sets used in the analyses in this thesis (except in the $\nu/\bar{\nu}$ separated analysis), including the interaction mode fractions estimated by MC, for all reconstructed sub-samples described in this section. We see that all selected e-like and μ-like sub-samples have a high purity of ν_e and ν_μ CC interactions, respectively.

In terms of energy coverage, looking back to Fig. 18 in Sec. 4.1, the MC true neutrino energy distributions of the sub-samples indicate that each has a different sensitive range. The FC μ-like sample extends above 10 GeV, while the e-like sample extends up to 100 GeV (indicating that μ-like events in the fiducial volume beyond 10 GeV tend to exit the inner detector and be classified as PC events). The

---

26 That is to say, the uncertainty of E_{rec} in the MC based on e.g. the PMT models, not the variance of E_ν.

| | NUMBER OF EVENTS | | | | | | | | INT. TYPE [%] | | |
	SK-I		SK-II		SK-III		SK-IV		TOTAL		ν_e CC	ν_μ CC	NC
FC e-LIKE													
sub-GeV single-ring	3288	(3104.7)	1745	(1632.8)	1209	(1100.7)	4251	(4072.8)	10493	(9911.0)	94.1	1.6	4.4
multi-GeV single-ring	856	(842.8)	396	(443.7)	274	(299.5)	1060	(1080.0)	2586	(2666.0)	86.3	4.9	8.8
multi-GeV multi-ring	449	(470.1)	267	(252.1)	140	(161.9)	634	(654.9)	1490	(1539.0)	73.0	10.9	16.1
FC μ-LIKE													
sub-GeV single-ring	3184	(3235.6)	1684	(1731.8)	1139	(1152.0)	4379	(4394.7)	10386	(10514.0)	0.9	94.2	4.9
multi-GeV single-ring	712	(795.4)	400	(423.9)	238	(273.9)	989	(1051.5)	2339	(2544.7)	0.4	99.5	0.1
multi-GeV multi-ring	603	(656.5)	337	(343.8)	228	(237.9)	863	(927.8)	2031	(2166.0)	3.4	91.1	5.5
PC													
stopping	143	(145.3)	77	(73.2)	54	(53.3)	237	(229.0)	511	(500.8)	12.7	82.7	4.6
through-going	759	(783.8)	350	(383.0)	290	(308.8)	1093	(1146.7)	2492	(2622.3)	0.8	98.6	0.6
UPMU													
stopping	432.0	(433.7)	206.4	(215.7)	193.7	(168.3)	492.7	(504.1)	1324.8	(1321.8)	1.0	98.7	0.3
non-showering	1564.4	(1352.4)	726.3	(697.5)	612.9	(504.1)	1960.7	(1690.3)	4864.3	(4244.4)	0.2	99.7	0.1
showering	271.7	(291.6)	110.1	(107.0)	110.0	(126.0)	350.1	(274.4)	841.9	(799.0)	0.1	99.8	0.1

Table 11: Event numbers for each SK period for each sub-sample. Numbers in brackets are the MC expectation, based on the HKKM11 [19] flux model, which are calculated separately for each SK period and then scaled to the data livetime. The oscillation and solar activity weights are included. UPMU has fractional data event numbers due to the subtraction of horizontally-going cosmic muon backgrounds. The relative frequency of interaction type in each sub-sample, estimated from the MC datasets averaging over all SK periods, are shown in the last three columns.
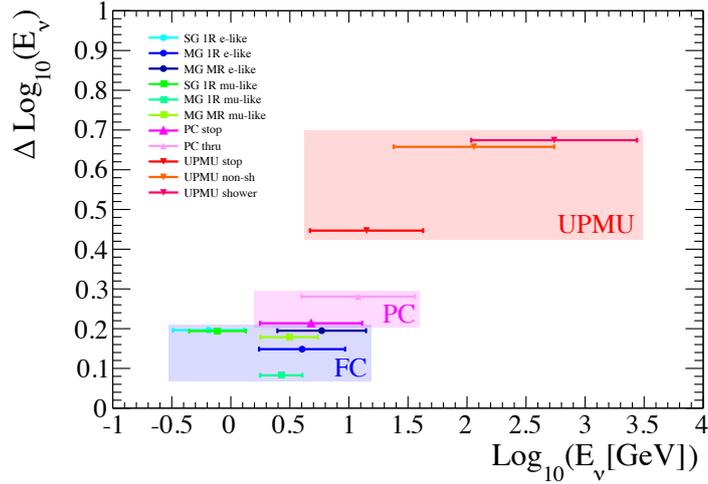
Figure 36: The energy resolution in the true neutrino energy E_ν for each sub-sample, estimated by MC events.

FC *e*-like sample also has a lower neutrino energy threshold than the μ-like sample (100 MeV as opposed to 200 MeV), due to the difference in Cherenkov threshold between $e^\pm$ and $\mu^\pm$. The PC and UPMU samples approximately range from 1 GeV to 10 TeV, split across several sub-samples which again provide some discrimination of the true neutrino energy range.

We show in Fig. 36 the energy resolution of each sub-sample, taken from the MC. We see that for the FC sample our estimation of the true neutrino energy per event is distributed around the true energy E_ν with a 1σ width of 10∼20% in $\log_{10}(E_\nu)$, but slightly less accurate for PC (20∼30%) and somewhat worse for UPMU (45∼70%) due to the missing energy deposited outside of the detector.

Part III

THE ANALYSES

I may not be very good at what I do, but at least I'm slow.

— FuzzDad

MEASUREMENT OF THE ENERGY SPECTRA OF THE NEUTRINO FLUX

5.1 INTRODUCTION

As explained in Sec. 1.1, knowledge of the atmospheric neutrino flux density is of vital importance to several current and future generation neutrino experiments, and to other rare event searches such as those for dark matter, proton decay, or astrophysical neutrinos. In particular, the flux as a function of neutrino energy should be well-known, as the neutrino oscillation effects depend strongly on the neutrino energy, and many rare event searches need to know the atmospheric neutrino background at particular energy scales.

In this section, we aim to reconstruct the energy spectrum of the $\nu_e + \bar{\nu}_e$ and $\nu_\mu + \bar{\nu}_\mu$ fluxes at the Super-Kamiokande detector, to a high precision that should be useful for testing to the current flux simulations, over a wide energy spectrum. After this reconstruction has been explained in detail, we further attempt to reconstruct the ν_e, $\bar{\nu}_e$, ν_μ, and $\bar{\nu}_\mu$ fluxes separately. The $\nu/\bar{\nu}$ separated samples have much smaller statistics and lower purity, such that the number of data points is expected to be smaller and the systematic errors larger; such a study can nonetheless add new information, which may help to understand e.g. any systematic bias seen in the $\nu_e + \bar{\nu}_e$ and $\nu_\mu + \bar{\nu}_\mu$ measured fluxes. Although theoretically the systematic errors on the predicted atmospheric $\nu/\bar{\nu}$ ratios are at most a few percent, this measurement is the first direct test of these predictions using atmospheric neutrino data.

Hereafter, until the section on the $\nu/\bar{\nu}$ separated measurement, antineutrinos are not mentioned explicitly and should be assumed to be included. Also, while ν_τ induced events have been observed at Super-Kamiokande with a high significance [130] compared to the no ν_τ hypothesis, the detector cannot yet make an accurate measurement of their flux, and they are not considered here.

Authorship note: this chapter (5) presents work performed in participation with Prof. Kimihiro Okumura (ICRR). With regards to the $\nu + \bar{\nu}$ study, while the author of this thesis made contributions such as calculating the systematic errors and performing cross-checks of the unfolding procedure, Prof. Okumura was the originator of the study and primary contributor to the analysis. The $\nu/\bar{\nu}$ separated analysis, on the other hand, is an original contribution of this author.

5.2   EVENT CLASSIFICATION

We use all sub-samples and events shown in Tab. 11. Noting that the FC sample may be well separated by its ν_e and ν_μ components, and that the PC and UPMU samples generally contain ν_μ interactions, we define two data samples for the energy spectrum measurement:

1. a ν_e sample containing FC e-like events,

2. a ν_μ sample from all FC μ-like, PC, and UPMU events,

which gives us a generally high-purity selection of ν_e or ν_μ induced CC interactions, across a wide energy range. For each sample, the backgrounds are defined as CC interactions of other flavour neutrinos, and NC interactions of all flavour neutrinos (non-neutrino backgrounds are assumed to be negligible).

In each sub-sample, events are binned by the reconstructed energy $E_{\rm rec}$, as defined in Sec. 4.2.2. Essentially for single-ring events, $E_{\rm rec}$ is calculated with the reconstructed lepton momentum, and for multi-ring events the linear sum of each particle's reconstructed momentum is taken. The binning is defined as shown in Tab. 12, in a unit of $\log_{10}\left(E_{\rm rec}[{\rm GeV}]\right)$. These bin widths were determined considering the energy resolutions of the true neutrino energy E_ν for each sub-sample, which are typically $\sim$ 20% in $\log_{10}\left(E_\nu[{\rm GeV}]\right)$ for the FC and PC stopping events, and larger at 40 $\sim$ 60% for the PC through-going and UPMU events (as the energy deposited outside the detector cannot be observed). No energy binning is performed for UPMU showering or non-showering events, since there is little separating power of the true neutrino energy from the observed energy. We refer hereafter to these defined bins simply as the "data bins", and the number of events in each data bin j as M_j. Figure 37 shows the values of M_j for all SK-I to SK-IV data.

5.3   FLUX UNFOLDING

5.3.1   *Overview*

To measure the atmospheric neutrino energy spectrum, we employ an unfolding method. This is a class of method where a true spectrum is deconvolved from an experimentally measured one, based on a knowledge of the experimental reconstruction. This class of methods is in contrast to forward-fitting methods, where the reconstructed spectrum is compared to model predictions that have also been passed through the simulated experimental reconstruction. A benefit of unfolding methods is that they allow direct comparisons of the unfolded spectra between different experiments, in a model-independent way. In our case, the reconstructed energy spectrum M_j

SUB-SAMPLE	BINS	j	BIN EDGES $\log_{10}\left(E_{rec}[\text{GeV}]\right)$
FC e-LIKE			
sub-GeV single-ring	5	1→5	-1.0, -0.6, -0.4, -0.2, 0.0, 0.2
multi-GeV single-ring	4	6→9	0.0, 0.4, 0.7, 1.0, 3.0
multi-GeV multi-ring	3	10→12	0.0, 0.4, 0.7, 3.0
FC μ-LIKE			
sub-GeV single-ring	5	13→17	-0.8, -0.6, -0.4, -0.2, 0.0, 0.2
multi-GeV single-ring	2	18→19	0.0, 0.4, 2.0
multi-ring	4	20→23	-1.0, 0.0, 0.4, 0.7, 2.0
PC			
stopping	2	24→25	-1.0, 0.4, 2.0
through-going	4	26→29	-1.0, 0.0, 0.4, 0.7, 3.0
UPMU			
stopping	3	30→32	0.0, 0.4, 0.7, 3.0
non-showering	1	33	–
showering	1	34	–

Table 12: Energy binning definitions for the data bins M_j for the 11 sub-samples.
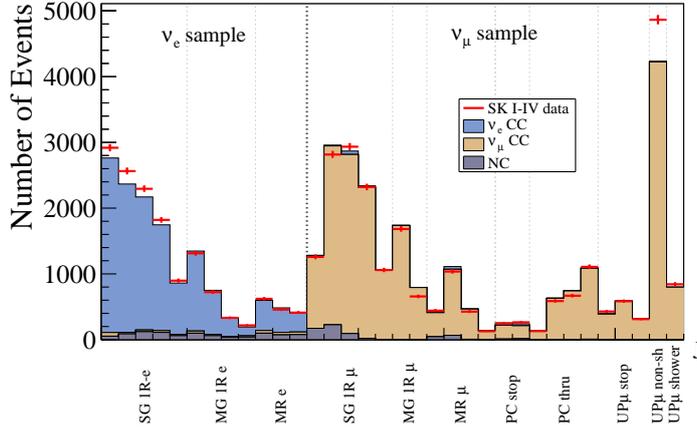
Figure 37: Number of events for each reconstructed energy bin M_j for each sample, for the entire SK-I to SK-IV period. The filled histograms are the MC truth predictions, and red represents the data counts with statistical error. Some inconsistency between the MC and data is seen, which can be accounted for within the estimated systematic errors (not plotted here).

(with binning defined in Tab. 12) will be unfolded into a true neutrino energy spectrum, of correct flavour CC interactions only, which we denote N_i^{CC} (with binning to be defined later, but note that the definition may differ from the reconstructed energy binning).

The relationship between the reconstructed data and the true data is in general expressed by the experimental response matrix A_{ji} as

$$M_j = A_{ji}N_i^{CC}, \tag{49}$$

where A_{ji} can be estimated by the detector MC, and accounts for the inability of the detector to reconstruct perfectly the true neutrino energy (e.g. energy smearing or inefficiency). We can write the inverse relationship using the unfolding matrix U_{ij} as

$$N_i^{CC} = U_{ij}M_j. \tag{50}$$

However, taking U_{ij} as A_{ji}^{-1} is in principle a poor approach, as the response matrix will not have been estimated perfectly (and may not even be invertible). Various more advanced techniques or algorithms are available, allowing for a more reliable unfolding to be performed [131].

5.3.2 *Bayesian unfolding*

We adopt an iterated Bayesian method [132], using the `RooUnfold` library [133] for the practical implementation. The method is known as Bayesian due to its use of Bayes' theorem in the construction of the unfolding matrix at each iterative step.[1] The `RooUnfold` library is chosen as it is known to be reliable and easy to implement.

The specific details of the unfolding procedure are as follows. The first estimation of the detector response matrix A_{ji} is made using the SK MC dataset, but using only correct flavour CC events reaching the final samples. The background events (i.e. events that were actually wrong flavour or NC interactions) are considered as an extra possible true cause bin i, and filled into an extra column of the matrix. The contribution of background events in the data can be seen in Tab. 11 to be small, and can be essentially subtracted during the unfolding procedure to leave a negligible impact on the final results. We also apply the normalization condition $\sum_j A_{ji} = 1$, which means that detector inefficiencies are not accounted for in the matrix at this stage – but will be accounted for later, when converting from the measured number of events to the actual neutrino flux values. Figure 38 shows the initial estimation of A_{ji} (not including the background events column).

To construct the unfolding matrix, we first define the notation $P(j|i)$ as the probability for an event in true energy bin i to be detected

---

[1] Although the final result is not strictly related to the input data in a Bayesian way [131].

in reconstructed energy bin j, which are exactly the elements of the response matrix A_{ji}. We then define initial prior probabilities $P_0(i)$ for a final sample event to fall in true energy bin i as

$$P_0(i) = \frac{N_{\text{MC},i}^{\text{CC}}}{\sum_k N_{\text{MC},k}^{\text{CC}}}, \tag{51}$$

which is simply the MC default true spectrum $N_{\text{MC},i}^{\text{CC}}$ normalized to one. We can then state Bayes' theorem

$$P(i|j) = \frac{P(j|i)P_0(i)}{P_0(j)}, \tag{52}$$

where the normalization constant $P_0(j)$ is calculated as

$$P_0(j) = \sum_i P(j|i)P_0(i). \tag{53}$$

Now that we have the estimated inverted probabilities $P(i|j)$, we can use our data M_j to make a *first estimation* of the true flux values $\hat{N}_i^{\text{CC}}$ as

$$\hat{N}_i^{\text{CC}} = \sum_j P(i|j)M_j. \tag{54}$$

As this estimation takes inputs from both MC and real data, it generally lies somewhere between the prior probabilities and the true number of events. Thus, it is useful to proceed in an *iterative* way, by inserting the normalized $\hat{N}_i^{\text{CC}}$ as an updated set of prior probabilities $P_0(i)$ in (52) to generate new $P(i|j)$, and in turn iteratively update $\hat{N}_i^{\text{CC}}$. The final iteration is denoted N_i^{CC}, where the operation of the above procedure on the data essentially took the place of the unfolding matrix U_{ij} in Eq. (50).

This method is seen to cause a variable degree of regularization (or smoothing) depending on the number of iterations; a low number of iterations may be too close to the statistically smooth prior values and not fully reflect the information input by the data, while a high number of iterations will lead statistical fluctuations in the data to distort the shape of the unfolded spectrum. In general any prior (including a flat prior) will thus cause some bias in the unfolded spectrum which cannot be perfectly corrected for, but importantly can be accurately estimated and included as a systematic error; this will be shown in Section 5.4. In principle, the prior distribution should contain our best knowledge of the true spectrum, and the bias is minimized by taking the nominal MC distribution as the initial prior. In practice, the iterative Bayesian method is strongly data-driven and produces excellent results compared to other unfolding methods after only a few iterations; we set the number of iterations to five based on MC optimizations.
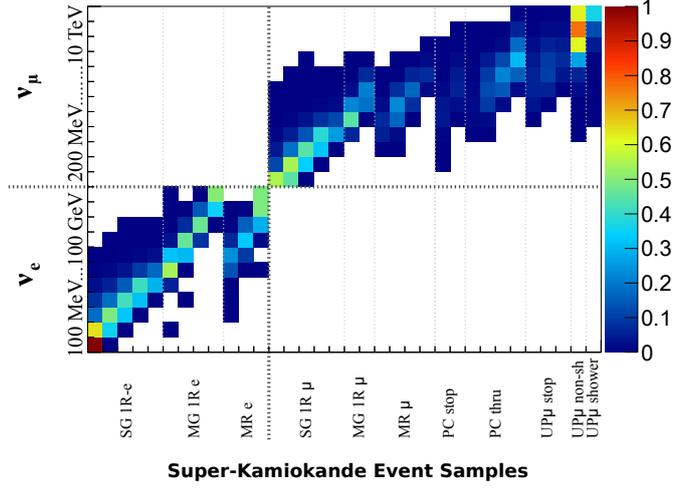
Figure 38: The initial estimation of the detector response matrix A_{ji} for the SK-I period, where the x-axis represents the M_j bins as defined in Tab. 12, and the y-axis the N_i^{CC} bin as defined in Tab. 13. (If we take the arrangement of the axis definitions seriously however, the figure actually shows A_{ij}^{-1}.)

5.3.3 *Binning definitions for the output spectra*

The energy range of the unfolded atmospheric neutrino flux spectrum N_i^{CC} is defined in units of $\log_{10}(E_\nu[\text{GeV}])$, for the ν_e (ν_μ) flux in the range -0.8→2.0 (-0.6→3.0), and divided into 11 (12) bins of variable size. The binning is shown later, together with the results, in Tab. 13 on page 105. This energy range, and the bin widths, are again determined considering the coverage and resolution of the true neutrino energy in the reconstructed data, as was shown in Fig. 36 on page 88. Smaller bins of width 0.2 are adopted below 10 GeV due to the relatively finer energy resolution of the FC sample, and wider bins are adopted for higher energies due to the deterioration of the energy resolution in the PC and UPMU samples.

5.3.4 *Calculation of the absolute flux values*

We obtain the measured neutrino flux values at the detector position $\Phi_i^{\nu_\alpha}$ (where $\alpha = e, \mu$) by using the predicted flux values $\Phi_{MC}^{\nu_\alpha}(\bar{E}_i)$, the predicted number of CC interactions $N_{MC,i}^{CC}$, and the unfolded number of CC interactions N_i^{CC}, as

$$\Phi_i^{\nu_\alpha} = \Phi_{MC}^{\nu_\alpha}(\bar{E}_i) \times \frac{\sum_{SK} N_{i,SK}^{CC}}{\sum_{SK} N_{MC,i,SK}^{CC}} \tag{55}$$

where $\sum_{SK}$ is a sum over the four SK periods, as considering the difference in detector configurations an unfolding must be performed

separately for each period (denoted by the subscript SK). Here, the predicted flux values $\Phi_{\text{MC}}^{\nu_\alpha}(\bar{E}_i)$ are calculated at the mean energy of each bin i, by integrating the predicted flux spectrum $\Phi_{\text{MC}}^{\nu_\alpha}(E)$ across the energy range of each bin (according to [134]) as

$$\Phi_{\text{MC}}^{\nu_\alpha}(\bar{E}_i) = \frac{1}{E_{i+1} - E_i} \int_{E_i}^{E_{i+1}} \Phi_{\text{MC}}^{\nu_\alpha}(E) dE \tag{56}$$

where E_i and E_{i+1} are the lower and upper edges of i-th energy bin. Here, the predicted flux spectrum is itself the integral of the differential atmospheric neutrino flux model by solid angle (including neutrino oscillation) as follows:

$$\Phi_{\text{MC}}^{\nu_\alpha}(E) = \int_{4\pi} d\Omega \left\{ \phi_{\text{MC}}^{\nu_\alpha}(E,\theta,\phi) \times O^{\nu_\alpha}(E,\theta) + \phi_{\text{MC}}^{\bar{\nu}_\alpha}(E,\theta,\phi) \times O^{\bar{\nu}_\alpha}(E,\theta) \right\},$$

$$\tag{57}$$

where $\phi_{\text{MC}}^{\nu_\alpha,\bar{\nu}_\alpha}(E,\theta,\phi)$ is the differential neutrino or anti-neutrino flux, as a function of neutrino energy E and the zenith and azimuthal angles of arrival direction θ and ϕ, which is taken by default from the HKKM11 model with the solar activity parameter set to its middle degree. Furthermore, $O^{\nu_\alpha,\bar{\nu}_\alpha}(E,\theta)$ is the neutrino or anti-neutrino oscillation probability from the production point to the detector; as discussed in the introduction, this is calculated with the standard three-flavor oscillation model with the oscillation parameters as shown in Table 1 (including the matter effect when propagating inside the Earth). Finally, for shorthand, we also define Φ_i (with no superscript) as the final binning containing both the $\Phi_i^{\nu_e}$ and $\Phi_i^{\nu_\mu}$ values.

5.4 STATISTICAL AND SYSTEMATIC UNCERTAINTIES

5.4.1 *Overview*

In addition to the statistical uncertainties, the systematic errors in the analysis can be considered to relate to one of four categories:

    A. event reconstruction by the detector,

    B. neutrino cross-sections,

    C. neutrino oscillation parameters,

    D. systematic bias in the unfolding procedure.

As will be described in this section, the uncertainties on A and B are estimated by the SK collaboration and are treated by modifying the data bins M_j. Next, the uncertainties on C are estimated by the PDG [59] and treated by modifying $O^{\nu_\alpha}(E,\theta)$ and $O^{\bar{\nu}_\alpha}(E,\theta)$ in Eq. (57).

Finally, the effect of D is estimated specifically for this analysis, and is also treated by modifying M_j.[2]

While it may seem more natural to think of the detector uncertainties A as affecting A_{ji} and the neutrino uncertainties B and C as affecting $N^{\text{CC}}_{\text{MC},i}$, our treatment of modifying M_j combined with the estimation of D should be equivalent, and much less computationally intensive. The equivalence was checked and confirmed with a few sample systematics.

5.4.2   *Reconstruction and cross-section errors*

The reconstruction and cross-section systematic error effects are calculated from the SK systematic database, which was described in Sec. 4.5.4. For this analysis all errors in the database, as summarized in Appendix A, are considered. Their effects are calculated as systematic error coefficients f_{jk}, which represent the fractional shift in each data bin M_j resulting from a 1σ shift of the k-th systematic error source.

For this analysis, the modified expectation of the number of events in each bin $\tilde{M}_j$ considering the effects of all errors would then become

$$\tilde{M}_j = M_j \times \left(1 + \sum_{k}^{N_{\text{sys}}} f_{jk} g_k \right), \tag{58}$$

where N_{sys} is the number of systematic error sources, and the g_k represents the applied strength of each systematic in units of σ. The estimated error sizes on the final measured flux values $\Phi_i^{\nu_\alpha}$ will be propagated by a toy MC method, which is described later in this section, using randomly-generated Gaussian sets of g_k.

5.4.3   *Oscillation parameter uncertainties*

The uncertainties of the oscillation parameters are shown in Tab. 1, and are assumed to be Gaussian (except for δ_{CP} where uniform probability is considered, and for the mass hierarchy where both possible values are considered). The correlation among these parameters are not taken into account. By directly shifting the values of $O^{\nu_\alpha}(E, \theta)$ and $O^{\bar{\nu}_\alpha}(E, \theta)$ in Eq. (57), the error on the final values $\Phi^{\nu_\alpha}_{MC}$ in Eq. (57) may be directly estimated. We generate toy flux spectra $\Phi^{\nu_\alpha,t}_{MC}$ with randomly shifted oscillation uncertainties.

---

2 We will see that although unfolding methods are sometimes disfavoured in the high-energy physics community (due to inaccurate error reporting in previous cases), if the particular case under study is well-understood then the errors can be accurately estimated. Then, the bias originating due to inaccuracies in the MC or theoretical model should be no worse than in commonly-used forward-fitting methods.

5.4.4 *Regularization uncertainties*

Another contribution to the uncertainty in the values N_i^{CC} is the regularization error; although the iterative procedure is data-driven, some influence from the MC shape still remains. In other words, the regularization error is due to inaccuracy in the initial estimation of the response matrix A_{ji} and Bayesian prior $N_{MC,i}^{CC}$, but must be applied in addition to the systematic errors on those estimations (which were applied directly on the data bins M_j). Any bias due to regularization should be relatively small, but it may be noticeable in unfolded bins with low statistics. This bias cannot be corrected for exactly, but we may estimate an associated error by unfolding pseudo-data sets with an energy spectrum reasonably far from the default MC prediction, and observing the resulting difference.

Such pseudo-data sets were produced from the MC data by re-weighting the true neutrino energy spectral index α and overall normalization γ, defining a new binned spectra $\Phi'_{MC,i}$ according to

$$\Phi'_{MC,i}(\bar{E}_i) = (1 + \Delta\alpha)\left(\frac{\bar{E}_i}{1\,\text{GeV}}\right)^{\Delta\gamma}\Phi_{MC,i}(\bar{E}_i). \tag{59}$$

A modification range of ± 0.05 in α and $\pm 20\%$ in γ was considered, which corresponds roughly to a 1σ deviation in the predicted spectra, according to the estimated systematic uncertainties in the flux model that were previously shown in Fig. 9 on page 25.

Figure 39 shows the comparisons of several input pseudo-data sets and their unfolded flux spectra, along with the fractional deviation between the input and unfolded output in each case, for both ν_e and ν_μ. The difference becomes largest in the highest energy bins, and the error sizes are $\pm 6\%$ and $\pm 8\%$ at most in the case of a ± 0.05 and $\pm 20\%$ change. The RMS of the differences are taken as the estimated uncertainties due to regularization, and implemented as a systematic error coefficient f_{jk}.

5.4.5 *Propagation of the statistical and systematic uncertainties*

For propagation of both the systematic and statistical uncertainties from M_j to the flux values Φ_i, we employ a toy MC method. [3] In this method, pseudo-data sets are generated by randomly fluctuating MC data according to their systematic and statistical error probability distribution functions (PDFs), as follows. First, the systematically shifted pseudo-data $\tilde{M}_j$ are calculated by Eq. (58) in Sec. 5.4.2, where the g_k are generated randomly each time, assuming that their PDFs

---

3 While RooUnfold provides an accurate theoretical calculation of the statistical errors [133], simultaneous treatment of systematic uncertainties within the software is not implemented.

(a) ν_e

(b) ν_μ
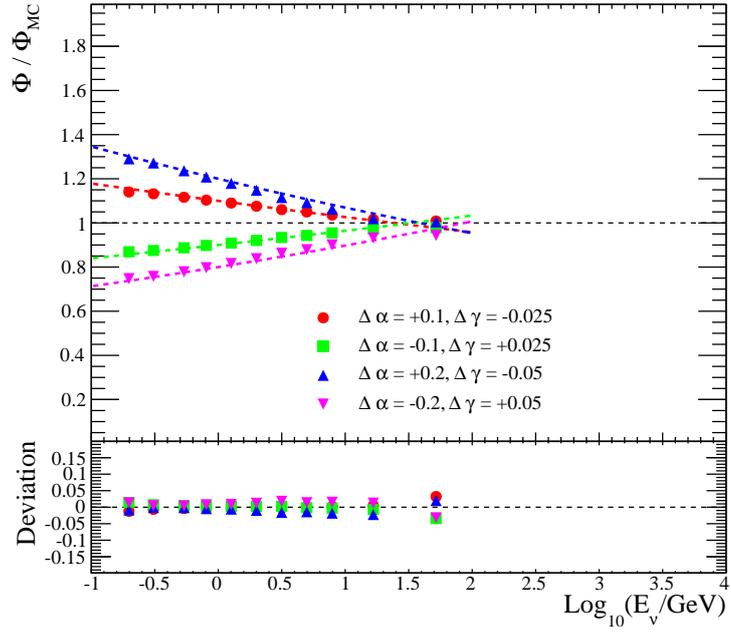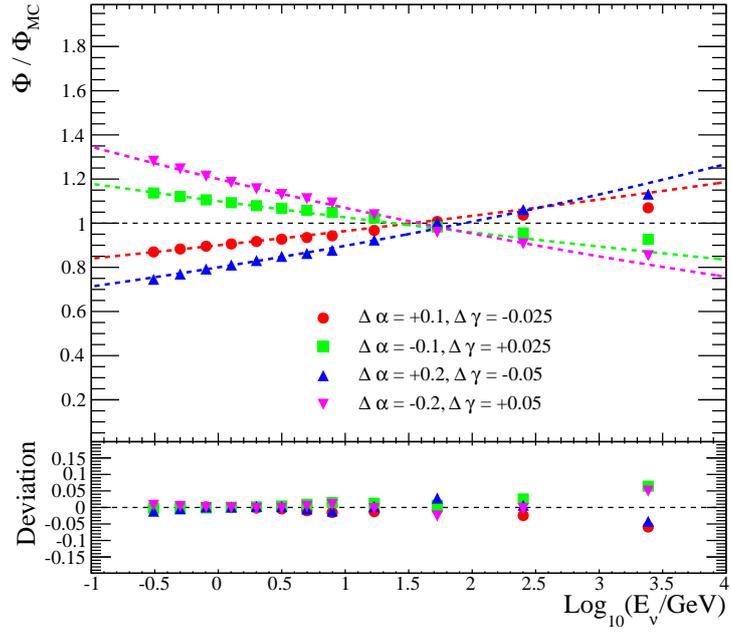
Figure 39: Pseudo-data input and the unfolded energy spectra (upper) and their ratios (lower) for ν_e and ν_μ. The vertical axis is the ratio to the HKKM11 [19] model Φ_{MC}. Four sets of pseudo-data inputs are tested, where $\Delta\alpha$ and $\Delta\gamma$ are the deviations of the flux normalization and spectral index respectively.

are Gaussian-distributed and uncorrelated. This includes the regularization error, as described in Sec. 5.4.4. Then, the statistical errors are applied by generating the final toy data M_j^t using Poisson distributions with means $\tilde{M}_j$. These data M_j^t are unfolded, using the method as described in the previous section, into toy measurements $N_i^{\mathrm{CC},t}$. Toy flux spectra $\Phi_{\mathrm{MC}}^{\nu_\alpha,t}(E)$ are also generated as was described in Sec. 5.4.3.

Two thousand sets of toy data $N_i^{\mathrm{CC},t}$ and $\Phi_{\mathrm{MC}}^{\nu_\alpha,t}(E)$ (for $t = 1...2000$) are generated, from which we calculate the set of toy flux values Φ_i^t by Eq. (55). The 1σ deviation of these toy measurements are taken as the final uncertainties $\Delta\Phi_i$, including all systematic and statistical error sources.

The error covariance matrix C_{ij}, which is a measure of the correlation of the effect of systematic errors in each pair of flux bins Φ_i, is also estimated at this stage using the toy MC sets as

$$C_{ij} = E\left[\left(\Phi_i^t - E[\Phi_i^t]\right)\left(\Phi_j^t - E[\Phi_j^t]\right)\right] \tag{60}$$

where E denotes the expected value, i.e. the mean, across the toy sets t.[4] The diagonally-normalized covariance matrix $\hat{C}_{ij}$, which will be used later for the χ^2 calculation, is defined as

$$\hat{C}_{ij} = C_{ij}/\Delta\Phi_i\Delta\Phi_j \tag{61}$$

and is shown in Fig. 40.

5.4.6 Error source comparisons

Figure 41 shows the breakdown of the estimated uncertainties for each error source. To produce this figure, the absolute errors of the five groups (statistical, detector response, neutrino interaction, neutrino oscillation, and regularization), are individually calculated, by repeating the above toy MC method but without inclusion of the other four groups. It can be seen that currently, the neutrino cross-section errors are the dominant source of systematic uncertainty.

5.5   RESULTS AND DISCUSSIONS

5.5.1   Measured flux spectra

Figure 42 shows the obtained ν_e and ν_μ energy spectra, using all SK-I to SK-IV data. The energy binning, average bin energy (x-axis values), and measured value and error (y-axis values and errors), are also written in Tab. 13.

---

4 Also note that in this case, the index j does not refer to the data bins, but is simply the second index of the covariance matrix.
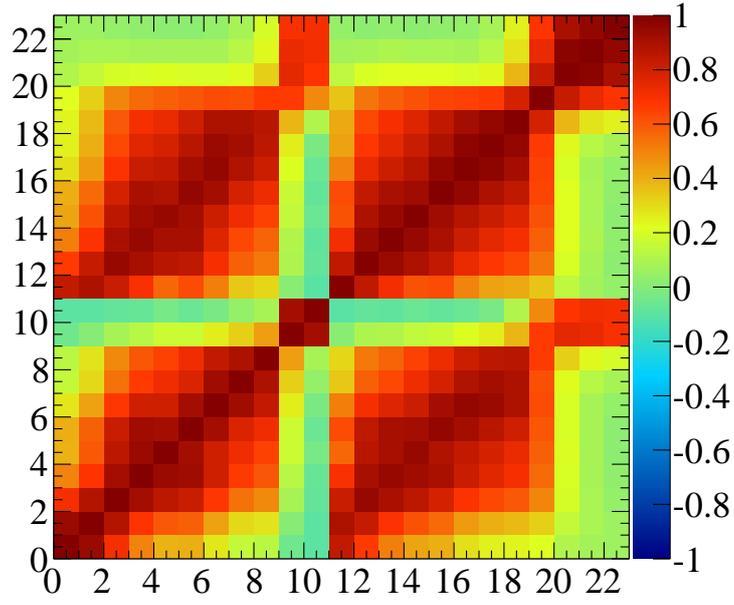
Figure 40: Normalized error correlation matrix $\hat{C}_{ij}$ of the observed flux values Φ_i. The matrix contains both ν_e and ν_μ, with ν_e ranging over bins $1 \rightarrow 11$ and ν_μ over bins $12 \rightarrow 23$. Strong correlations are seen between ν_e and ν_μ as many systematics apply to both flavours in similar ways, although some anticorrelated systematics also exist.
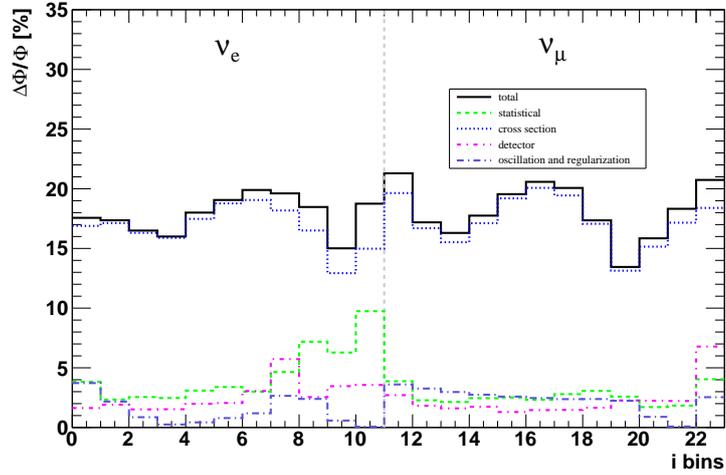
Figure 41: Breakdown of the estimated uncertainties in the measured flux values depending on each error group (statistical, detector response, neutrino cross section, neutrino oscillation, and regularization). The x-axis represents the output spectrum bins as defined in Tab. 13. The "total" histogram represents the error when all error sources are considered simultaneously.
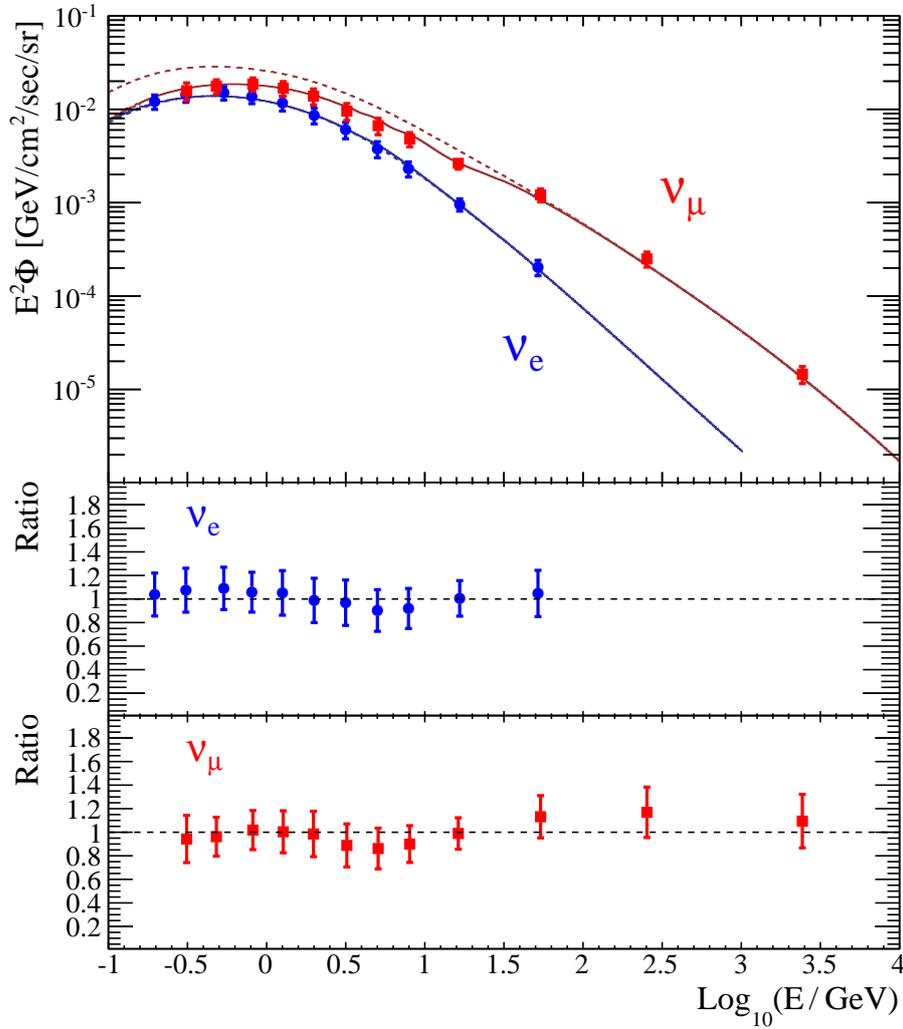
Figure 42: The measured flux energy spectra, using all SK-I to SK-IV data, for ν_e (blue) and ν_μ (red). The flux values are shown multiplied by E^2, following a common convention. Error bars include all statistical and systematic uncertainties. The curves show the HKKM11 [19] model's prediction with (solid) and without (dashed) neutrino oscillation (same color scheme). The lower part of the figure shows the ratio of data / HKKM11 prediction.

The measured energy spectrum agrees with oscillated HKKM11 flux, within the estimated uncertainties. The unoscillated flux is also plotted, where the deficit of ν_μ flux due to neutrino oscillation becomes apparent below 100 GeV.

5.5.2 *Comparison to flux models*
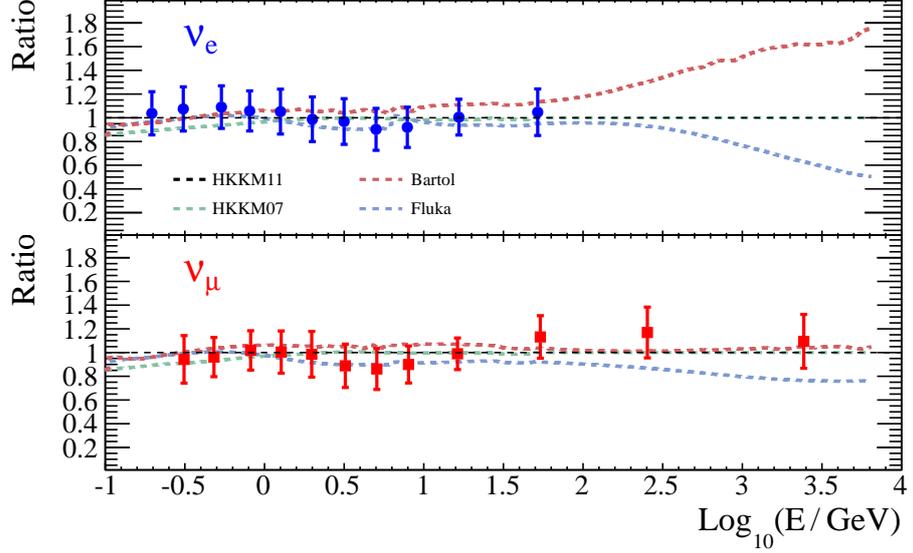
Figure 43: The measured energy spectra compared with several flux model predictions (dotted lines), HKKM11 [19], HKKM07 [18], Bartol [21], and Fluka [20]. Vertical axis is the ratio to the HKKM11 model. Error bars include all statistical and systematic uncertainties.

The observed fluxes are compared graphically to several flux models, including the HKKM11 [19], HKKM07 [18], Fluka [20], and Bartol [21] models in Figure 43.

To perform a quantitative comparison, we perform a χ^2 test which takes into account the error correlation matrix, defined as

$$\chi^2 = \sum_i^N \sum_j^N (\Phi_i - \Phi_{\text{MC},i})^T C_{ij}^{-1} (\Phi_j - \Phi_{\text{MC},j}) \tag{62}$$

where Φ_i is the observed flux and $\Phi_{\text{MC},i}$ is the expectation of the flux model being tested, at the i-th energy bin, and C_{ij} is the error covariance matrix (calculated as described in Sec. 5.4 and shown in Fig. 40). Both statistical and systematic uncertainties are thus taken into account in the χ^2 calculation, and the results are shown in Table 14 for three cases: ν_e and ν_μ, ν_e only, and ν_μ only. Considering the combined $\nu_e + \nu_\mu$ test, the χ^2 values do not differ significantly between the flux models, however HKKM11 has the best fit to the data. As our measurement prefers HKKM11 above HKKM07, we can identify

ENERGY RANGE $\log_{10}\left(E_i\,[\text{GeV}]\right)$	AVERAGE ENERGY $\log_{10}\left(\overline{E}_i\,[\text{GeV}]\right)$	MEASURED FLUX [GeV cm^{-2} sec^{-1} sr^{-1}]	ERROR [%]
		ν_e	
$-0.8 \rightarrow -0.6$	-0.71	1.21×10^{-2}	± 17
$-0.6 \rightarrow -0.4$	-0.51	1.46×10^{-2}	± 16
$-0.4 \rightarrow -0.2$	-0.27	1.50×10^{-2}	± 15
$-0.2 \rightarrow 0.0$	-0.09	1.37×10^{-2}	± 15
$0.0 \rightarrow 0.2$	0.10	1.16×10^{-2}	± 17
$0.2 \rightarrow 0.4$	0.30	8.55×10^{-3}	± 19
$0.4 \rightarrow 0.6$	0.50	6.09×10^{-3}	± 21
$0.6 \rightarrow 0.8$	0.70	3.73×10^{-3}	± 22
$0.8 \rightarrow 1.0$	0.90	2.32×10^{-3}	± 20
$1.0 \rightarrow 1.5$	1.22	9.42×10^{-4}	± 15
$1.5 \rightarrow 2.0$	1.72	2.03×10^{-4}	± 18
		ν_μ	
$-0.6 \rightarrow -0.4$	-0.51	1.58×10^{-2}	± 23
$-0.4 \rightarrow -0.2$	-0.32	1.77×10^{-2}	± 18
$-0.2 \rightarrow 0.0$	-0.09	1.86×10^{-2}	± 16
$0.0 \rightarrow 0.2$	0.10	1.68×10^{-2}	± 18
$0.2 \rightarrow 0.4$	0.30	1.38×10^{-2}	± 20
$0.4 \rightarrow 0.6$	0.51	9.59×10^{-3}	± 23
$0.6 \rightarrow 0.8$	0.71	6.68×10^{-3}	± 23
$0.8 \rightarrow 1.0$	0.90	4.79×10^{-3}	± 19
$1.0 \rightarrow 1.5$	1.21	2.62×10^{-3}	± 14
$1.5 \rightarrow 2.0$	1.73	1.20×10^{-3}	± 14
$2.0 \rightarrow 3.0$	2.40	2.49×10^{-4}	± 16
$3.0 \rightarrow 4.0$	3.39	1.46×10^{-5}	± 19

Table 13: Neutrino flux binning N_i^{CC}, and measurement results, using SK-I to SK-IV data. The measured flux values Φ_i are given multiplied by $\overline{E}_i^2$, following a common convention. The error includes both statistical and systematic uncertainties.

	χ^2		
	ν_e (11 DOF)	ν_μ (12 DOF)	ν_e AND ν_μ (23 DOF)
HKKM11	8.5	19.5	21.8
HKKM07	15.4	21.7	22.2
BARTOL	10.9	28.8	30.7
FLUKA	9.0	18.7	25.6

Table 14: The results of a χ^2 test of several flux models against the measured flux, according to Eq. (62). The number of degrees of freedom (DOF) in each test are also shown.

the updates to the hadron simulation [19] (which cause changes in the energy region below 1 GeV) as bringing slightly better agreement with the data. Looking at each flavour individually, the agreement for the predicted ν_e spectra for each model is good, however the ν_μ test shows a slight inconsistency at the 2.0 σ level for the HKKM07 model, and at the 3.0 σ level for the Bartol model.[5]

We also evaluate the agreement of the data with each model's normalization γ and spectral index α, by testing modified energy spectra $\Phi'_{\mathrm{MC},i}$ based on each model $\Phi_{\mathrm{MC},i}$, which were defined previously in Eq. (59). The χ^2 is then calculated by Eq. (62), and the best-fit values of the $\Delta\alpha$ and $\Delta\gamma$ parameters are obtained by minimizing the χ^2, and are shown in Fig. 44. The normalization and spectral index agrees within the 1 σ error for every model, except from the fitted spectral index of the Fluka ν_μ spectrum which deviates by 2.4 σ.

In summary, while none of the current flux models are generally inconsistent with our data measurement, there are some tensions in the ν_μ spectrum for the HKKM07 and Bartol models, and in the spectral index of the Fluka model. The HKKM11 model gives the best fit to the current data.

5.5.3 *Cross-check by SK period*

Considering the final unfolded spectra in Fig. 43, some systematic shift in shape is clearly seen for both the ν_e and ν_μ unfolded fluxes. In order to verify the consistency of this tendency across each SK period, the energy spectrum unfolding of each period is calculated and compared in Fig. 45, where only statistical errors are shown. While the error bars are relatively large, the spectral distortion seems consistent across SK periods, and not attributable to statistical error. The origin

---

5 Due to the error correlation, the χ^2 statistic for the $\nu_e + \nu_\mu$ test is not simply a combination of the χ^2 statistic of the individual tests. In principle we should also define a trials factor for the individual tests, since we are applying a related test to a subset of the data, but such a factor is assumed to be small as the tests are not completely correlated.
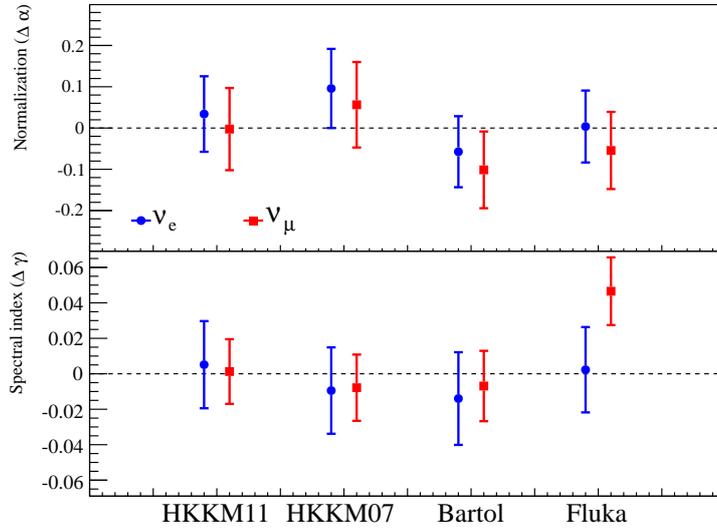
Figure 44: The best-fit and 1σ error of the normalization and spectra index parameters $\Delta\alpha$ and $\Delta\gamma$, obtained by χ^2 calculation against each model with the parameters artificially shifted.
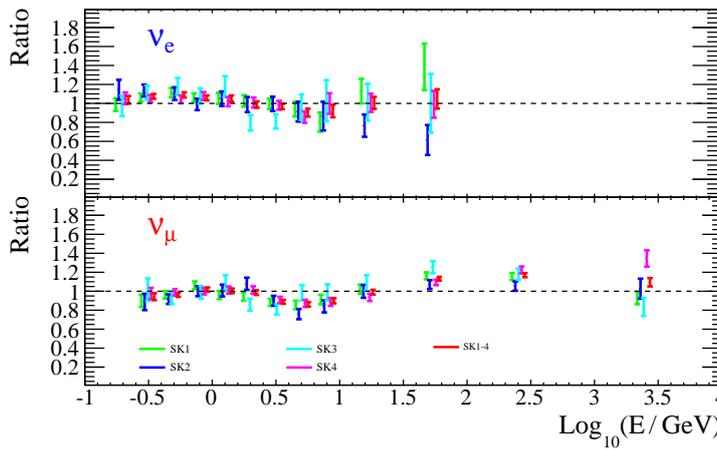
Figure 45: Flux ratio to the HKKM11 model for each SK period. Error bars include only statistical uncertainty in this figure.

of the spectral distortion is however within the range of the estimated systematic effects, in particular the neutrino interaction uncertainties, and so at present we cannot identify whether or not this indicates a direct disagreement with the HKKM11 model.

5.5.4 *Comparison to other measurements*
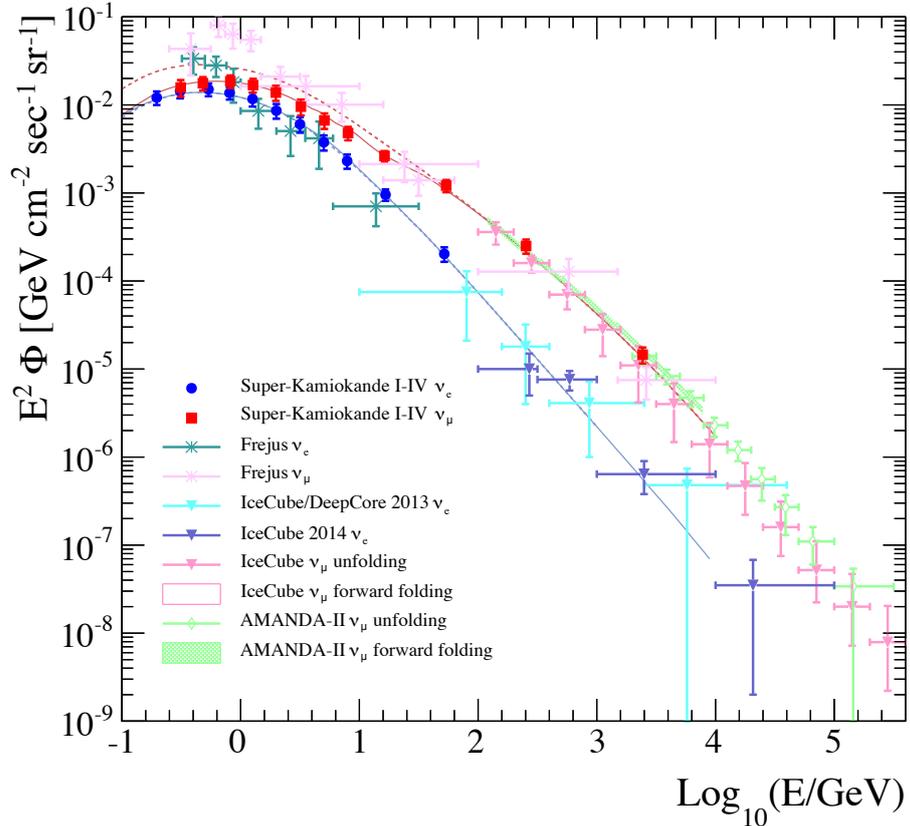
Figure 46: The measured energy spectra of atmospheric ν_e and ν_μ at SK, shown in comparison to measurements by Frejus [28], AMANDA-II [29, 30], and IceCube [31, 33, 32]. The HKKM11 flux curves are also shown in solid (with oscillation) and dashed (without oscillation) lines.

Figure 46 shows our flux measurement compared with the results from other experiments. Our measured data provides significantly improved precision up to 100 GeV, and a direct measurement is made for the first time below 320 MeV. At higher energies, our measurement overlaps with the AMANDA-II and IceCube ν_μ measurements. In the future, this overlap should allow a better constraint to be placed on the flux normalization at the energies beyond 100 GeV, which will be helpful to the understanding of the astrophysical neutrino flux.

5.6 MEASUREMENT OF NEUTRINO AND ANTINEUTRINO FLUXES

5.6.1 *Data selection and binning*

In this section we make use of the neutrino and antineutrino samples, as defined in Sec. 4.2.3. As a reminder, we cannot separate on an event by event basis, but have some ability to statistically "enrich" a sample's ν or $\bar{\nu}$ content by a certain percentage, based on the kinematic differences. In principle, we could then measure the $\nu/\bar{\nu}$ ratio directly if we completely understand and account for the cross-section and kinematic differences between any two data samples, although in practice this is quite difficult.[6] In the samples defined in Sec. 4.2.3 we thus additionaly made use of the theoretical $\nu/\bar{\nu}$ ratios in order to estimate the ν and $\bar{\nu}$ content in each sample – therefore we should keep in mind that any measurement with these samples is model-dependent. Nonetheless, the errors on the theoretical prediction of the $\nu/\bar{\nu}$ ratios are of the order 1%, and are not calculated directly from the normalizations, which have errors of the order 10%. A measurement of these samples can thus provide new information on the ν and $\bar{\nu}$ normalizations due to both kinematic or flux normalization differences.

The purity and statistics of the various $\nu/\bar{\nu}$ separated samples, as was shown in Tab. 6 on page 72, is quite variable. We wish to perform a sub-selection of these samples to define ν_e-like, $\bar{\nu}_e$-like, ν_μ-like, and $\bar{\nu}_\mu$-like samples, which are as high-purity[7] and high statistics as possible. With regards to the purity and statistics, increasing one generally comes at a cost to the other, so the criteria for choosing the "best" samples is somewhat arbitrary. We choose a guiding principle that a sample should be at least 70% pure to be included in the ν-like samples, and $\sim$ 30% ($\sim$ 40%) pure to be included in a $\bar{\nu}_e$-like ($\bar{\nu}_\mu$-like) samples. The requirements for the $\bar{\nu}$ samples are of course lower, due to the predominance of ν over $\bar{\nu}$ in the atmospheric flux before separation.

Firstly, both of the π^0-like samples and the "multi-ring other" sample are discarded, due to their high NC and ν_τ backgrounds. The PC and UPMU samples, while not shown in Tab. 6, generally have ν_μ and $\bar{\nu}_\mu$ purities of around 60% and 30% respectively, and these

---

6 For example, a measurement of the $\nu_\mu/\bar{\nu}_\mu$ ratio was attempted by considering the difference in the μ-like single-ring $n_d = 0$ and $n_d = 1$ samples (i.e. without or with a Michel electron), whereby about 20% of μ^- are captured on oxygen, thus enriching the ν_μ content of the $n_d = 0$ sample by a known amount. However, even in this simple case, many second-order kinematic effects were seen contribute, and the study was not completed on time to be included here.

7 While the estimated wrong flavour, wrong sign, and NC backgrounds will be subtracted during the unfolding process, and the uncertainty of the background subtraction is accounted for in the systematic error estimations, a high-purity sample minimizes the bias and also protects against spurious results in the (pessimistic) case where some systematics were not considered or were underestimated.

samples are also discarded. For the remaining FC samples, a study is performed examining the purity as a function of the reconstructed momentum E_{rec}, in case it may be possible to fulfill our stated purity requirements by choosing binning that uses only energy sub-ranges of some samples. This is found to be possible only for the "single-ring sub-GeV μ-like $n_d = 1$" sample, where we can raise the $\bar{\nu}_\mu$ purity to $\sim 40\%$ by using only events with $1000 < E_{\text{rec}} < 1330$ MeV.

The remaining samples and energy-ranges that satisfy our purity requirements are then selected, with the exception of the "single-ring sub-GeV e-like $n_d = 0$" sample – while the ν_e purity is high in this sample, the $\bar{\nu}_e$ contamination is somewhat higher than the other ν_e-like samples, and by discarding this sample we have the ability to perform an extremely good separation of ν_e and $\bar{\nu}_e$ in our ν_e-like sample. The final data samples and binning are shown in Tab. 15. Compared to the $\nu + \bar{\nu}$ study, the low statistics present more of a problem than the energy resolution, and the binning is generally chosen to keep statistical errors below 20% for each bin (even for the smaller SK-II dataset). Also, the maximum E_{rec} limit in some samples is also reduced, due to the much lower statistics at high energies. In summary, we obtain the following (relatively speaking):

- ν_e-like sample: high purity, low statistics

- $\bar{\nu}_e$-like sample: low purity, medium statistics (covers $E_{\text{rec}} > 1330$ MeV only)

- ν_μ-like sample: high purity, high statistics

- $\bar{\nu}_\mu$-like sample: medium purity, medium statistics (covers $E_{\text{rec}} > 1000$ MeV only).

For the output binning, shown later in Tab. 16, we again consider the predicted resolution and energy coverage of E_ν compared with E_{rec}, and modify the E_ν bins to minimize the statistical error in each bin. The first estimation of the detector response matrix by MC is shown in Fig. 47.

5.6.2   *Systematic errors*

5.6.2.1   *New systematic errors*

The accuracy of the division of samples into ν and $\bar{\nu}$-like have associated systematic errors that must be estimated, primarily coming from the accuracy of the neutron and Michel-electron tagging procedures, while other systematics were shown in [84] to be relatively smaller. These errors control migration between ν_e and $\bar{\nu}_e$, and between ν_μ and $\bar{\nu}_\mu$; we assume that the uncertainty on the background of wrong-flavour and NC interactions, which was already estimated
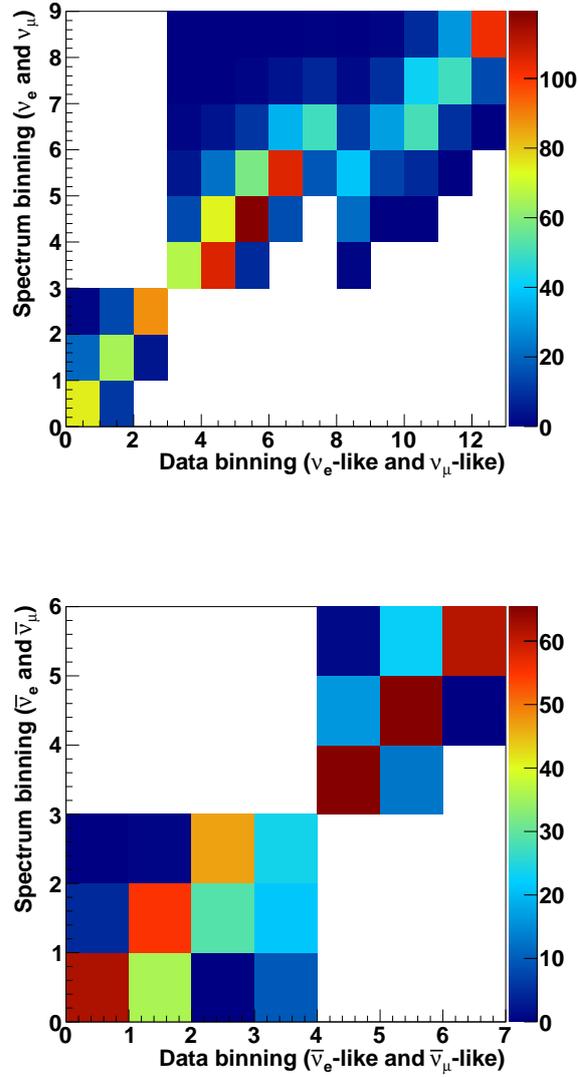
Figure 47: The initial estimation of the detector response matrix A_{ji}^{-1} for the $\nu/\bar{\nu}$ separated unfolding, showing ν (top) and $\bar{\nu}$ (bottom) for SK-I.

SUB-SAMPLE	BINS	BIN EDGES $\log_{10}\left(E_{\text{rec}}[\text{GeV}]\right)$
ν_e-LIKE		
single-ring sub-GeV e-like $n_d = 1$	2	-1.0, 0.3, 0.2
single-ring multi-GeV ν_e-like	1	0.0, 1.7
$\bar{\nu}_e$-LIKE		
single-ring multi-GeV $\bar{\nu}_e$-like	3	0.0, 0.25, 0.5, 2.2
multi-ring $\bar{\nu}_e$-like	1	0.0, 2.2
ν_μ-LIKE		
single-ring sub-GeV μ-like $n_d = 0$	5	-0.8, -0.6, -0.4, -0.2, 0.0, 0.2
single-ring sub-GeV μ-like $n_d = 2$	2	-0.8, -0.2, 0.2
multi-ring ν_μ-like	3	-0.8, 0.1, 0.3, 1.2
$\bar{\nu}_\mu$-LIKE		
single-ring sub-GeV μ-like $n_d = 1$	1	0.0, 0.2
single-ring multi-GeV $\bar{\nu}_\mu$-like	2	0.0, 0.4, 1.0

Table 15: Energy binning definitions for the data bins for the $\nu/\bar{\nu}$ separated unfolding.

in the $\nu_e + \bar{\nu}_e$ and $\nu_\mu + \bar{\nu}_\mu$ samples, is similar in the $\nu/\bar{\nu}$ separated samples. The decay electron tagging efficiency uncertainty is modeled in the SK systematic error database, causing normalization changes at around the 1~10% level depending on the sub-sample and SK period, and was included on the samples separated by n_d.

The uncertainty on the neutron tagging efficiency was estimated in [84] using a Americium/Beryllium neutron source, producing an accurately known intensity of low energy neutrons, placed in the tank; the results from this measurement were compared with the neutron tagging MC, and the difference between the two of $\sim 10\%$ was taken as the systematic error. For the MC itself, two simulations were run using different software packages (FLUKA and GEANT3), however the difference between them was negligible. The uncertainty on the decay electron tagging efficiency is included in the SK database, similarly being calculated by a comparison between MC and cosmic-ray muon decay electron data, and is at the 1% level. Finally, the systematics on the hadronic differences in the interactions were also estimated in [84] by comparing various theoretical (e.g. NEUT) and experimental data (e.g. the CHORUS ν_μ beam experiment), and found at the $2 \sim 10\%$ level, depending on the sample and energy range.

The final reconstruction systematic errors on the $\nu/\bar{\nu}$ separation for the subsamples used in this analysis were calculated as follows: approximately 5% for single-ring multi-GeV ν_e-like, 1% for single-ring multi-GeV $\bar{\nu}_e$-like, 1% for multi-ring $\bar{\nu}_e$-like, 1.5% for multi-ring ν_μ-
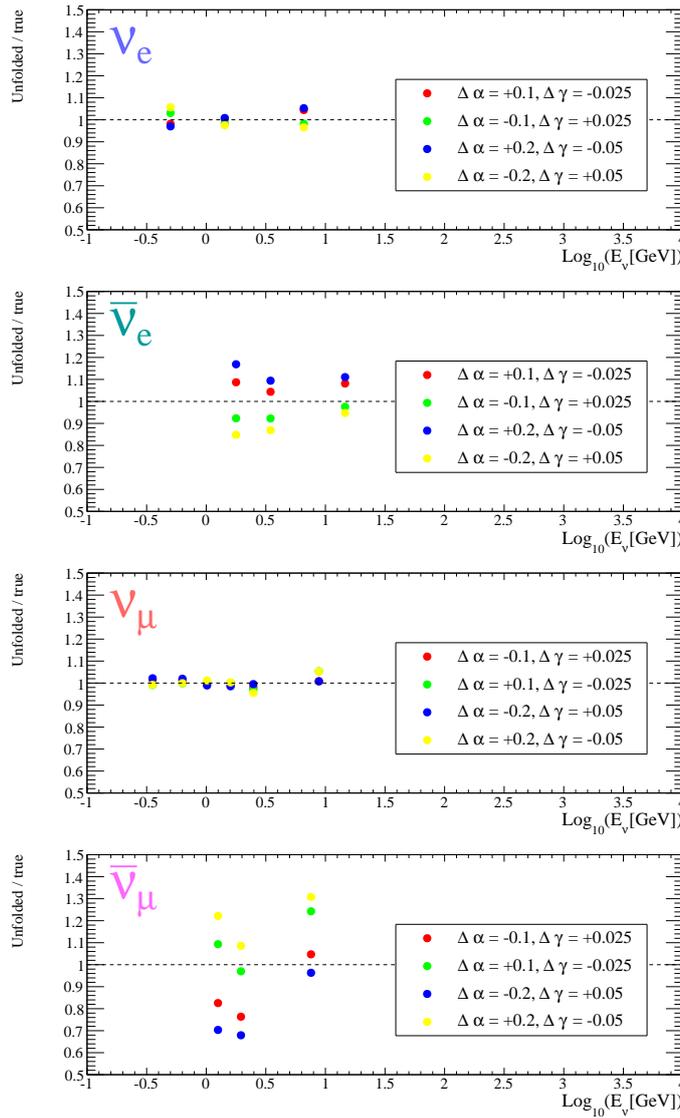
Figure 48: The regularization error estimated on the four $\nu/\bar{\nu}$ separated samples.

like, and 2% for single-ring multi-GeV $\bar{\nu}_\mu$-like, where the percentages are given by averaging the systematic error over the included energy ranges.[8]

5.6.2.2 *Recalculation of previously defined errors*

The regularization error, as was estimated in Sec. 5.4.4, was re-estimated using the same method, with the result shown in Fig. 48. For ν_e and ν_μ, generally the uncertainties are at the 1% level, rising to the 3% level for bins on the edge of the distributions; by keeping the number of output bins small, we see that we were able to retain a similar

---

8 Although the errors consider migration between e.g. ν_e and $\bar{\nu}_e$, the percentages may be larger on one sample than the other, due to the relative size of the samples.

accuracy to the $\nu + \bar{\nu}$ study, which had $< 1\%$ in the central bins and similarly 3% for the edge bins. For the antineutrino measurement however, the errors are much larger at the $5 \sim 15\%$ level. This is due to the fact that number of input bins is small, and they are crowded together a relatively narrow range of E_{rec}, which leads to a poorer resolution in true neutrino energy E_ν. Unfortunately, it is difficult to improve the choice of binning without a significant loss of purity.

The oscillation errors, SK systematic errors, and statistical errors were also re-calculated for the new input and output binnings. The oscillation errors are similar to the previous study, at the $2 \sim 3\%$ level for most output bins (in the previous study a small number of the $\nu_e + \bar{\nu}_e$ bins had an error of $< 1\%$, which was possible due to their narrow energy range). The systematic errors are also similar, with an additional $1 \sim 3\%$ error coming from the $\nu/\bar{\nu}$ separation uncertainties. The statistical errors on the other hand are somewhat higher in the new study, at the $8 \sim 10\%$ level for all ν_e, $\bar{\nu}_e$, and $\bar{\nu}_\mu$ bins, and slightly lower at 5% for ν_μ (whereas in the $\nu + \bar{\nu}$ study, the errors were generally at the $\sim 3\%$ level, although somewhat more energy-dependent and rising to $8 \sim 10\%$ in the very highest energy bins).

5.6.3   *Results and discussions*

The unfolded fluxes are shown graphically in Fig. 49, numerically in Tab. 16, and compared with various flux models in Fig. 50. The statistical error components are also plotted separately, to illustrate their relatively larger size compared to the previous study. The results mostly agree with the flux models within statistical and systematic error, except for the ν_e flux in the region below 1 GeV, which is measured higher than predicted, and the ν_μ flux in the 2 to 3 GeV region, which is slightly lower than predicted.

For the ν_e flux, considering the strong correlation of the systematic errors between energy bins, and that all ν_e data bins are higher than predicted, the ν_e measurement suggests that the flux is systematically higher than the default prediction. As the $\bar{\nu}_e$ flux is slightly lower, some contribution of this increase may be due to the $\nu_e/\bar{\nu}_e$ anticorrelated systematics; but as they are only at the few-percent level, there may also be some contribution due to cross-section differences (which have the largest systematic uncertainty, and affect the ν and $\bar{\nu}$ fluxes differently). For the data point below 1 GeV, we unfortunately do not have any $\bar{\nu}_e$ data to compare with, as we were not able to make a clear $\bar{\nu}_e$-enhanced measurement in that region. However if the ν_e flux is already systematically high, the upwards fluctuation of this bin (already significant at less than $2\,\sigma$) could also be reasonably explained as a statistical fluctuation. In any case, as the flux models agree well in

this region, this datapoint does not have strong discriminating power between the flux models.

For the ν_μ flux, we see the same general systematic shape as was seen in the $\nu_\mu + \bar{\nu}_\mu$ unfolded flux. However, the dip at around the $2 \sim 3$ GeV region is more pronounced in the unfolded ν_μ flux, and less so in the unfolded $\bar{\nu}_\mu$ flux. In this case the disagreement is hard to explain within the estimated $\nu_\mu / \bar{\nu}_\mu$ anticorrelated systematics, again at the few-percent level. Therefore, we again conclude that the default cross-sections may be more accurately modeled for $\bar{\nu}_\mu$ than ν_μ at this energy range.

Approaching the TeV scale, the errors on the predicted $\nu / \bar{\nu}$ ratios increase substantially (from the SK database, a 16% error at 1 TeV is assumed for the HKKM11 model). As can be extrapolated from Fig. 50, the $\nu_\mu / \bar{\nu}_\mu$ ratio of the Bartol flux strongly disagrees with the other models. We thus see that if they could be extended to a higher energy scale, this kind of measurement could be particularly useful in discriminating between the flux models.

In general, we presented here the first separated measurement of the ν and $\bar{\nu}$ atmospheric neutrino fluxes. Despite being a model-dependant measurement, it may be interpreted as providing more information on the relative normalizations of the ν and $\bar{\nu}$ fluxes when compared to the $\nu + \bar{\nu}$ measurement. Our results generally agree with the flux models, but do hint at some interesting ranges for further examination. In particular, the planned addition of gadolinium to SK should allow an excellent separation of neutrinos and antineutrinos, in turn allowing a more accurate investigation of the systematic differences in the event numbers, and a direct measurement of the $\nu / \bar{\nu}$ ratios.

ENERGY RANGE $\log_{10}\left(E_i \, [\text{GeV}]\right)$	AVERAGE ENERGY $\log_{10}\left(\overline{E}_i \, [\text{GeV}]\right)$	MEASURED FLUX $[\text{GeV cm}^{-2} \, \text{sec}^{-1} \, \text{sr}^{-1}]$	ERROR [%]
		ν_e	
-0.6→0.0	-0.3	1.09×10^{-2}	±20
0.0→0.3	0.16	6.77×10^{-3}	±21
0.3→1.5	0.82	2.15×10^{-3}	±20
		$\bar{\nu}_e$	
0.1→0.4	0.25	5.06×10^{-3}	±24
0.4→0.7	0.54	2.94×10^{-3}	±21
0.7→2.0	1.17	5.84×10^{-4}	±23
		ν_μ	
-0.6→-0.3	-0.45	8.27×10^{-3}	±18
-0.3→-0.1	-0.20	1.00×10^{-2}	±18
-0.1→0.1	0.01	1.00×10^{-2}	±17
0.1→0.3	0.20	8.41×10^{-3}	±18
0.3→0.5	0.40	5.08×10^{-3}	±23
0.5→1.5	0.95	2.23×10^{-3}	±22
		$\bar{\nu}_\mu$	
0.0→0.2	0.01	9.10×10^{-3}	±23
0.2→0.4	0.29	6.76×10^{-3}	±21
0.4→1.5	0.88	2.57×10^{-3}	±28

Table 16: Neutrino flux binning, and measurement results, for the $\nu/\bar{\nu}$ separated unfolding using SK-I to SK-IV data. The measured flux values Φ_i are given multiplied by $\overline{E}_i^2$, following a common convention. The error includes all statistical and systematic uncertainties.
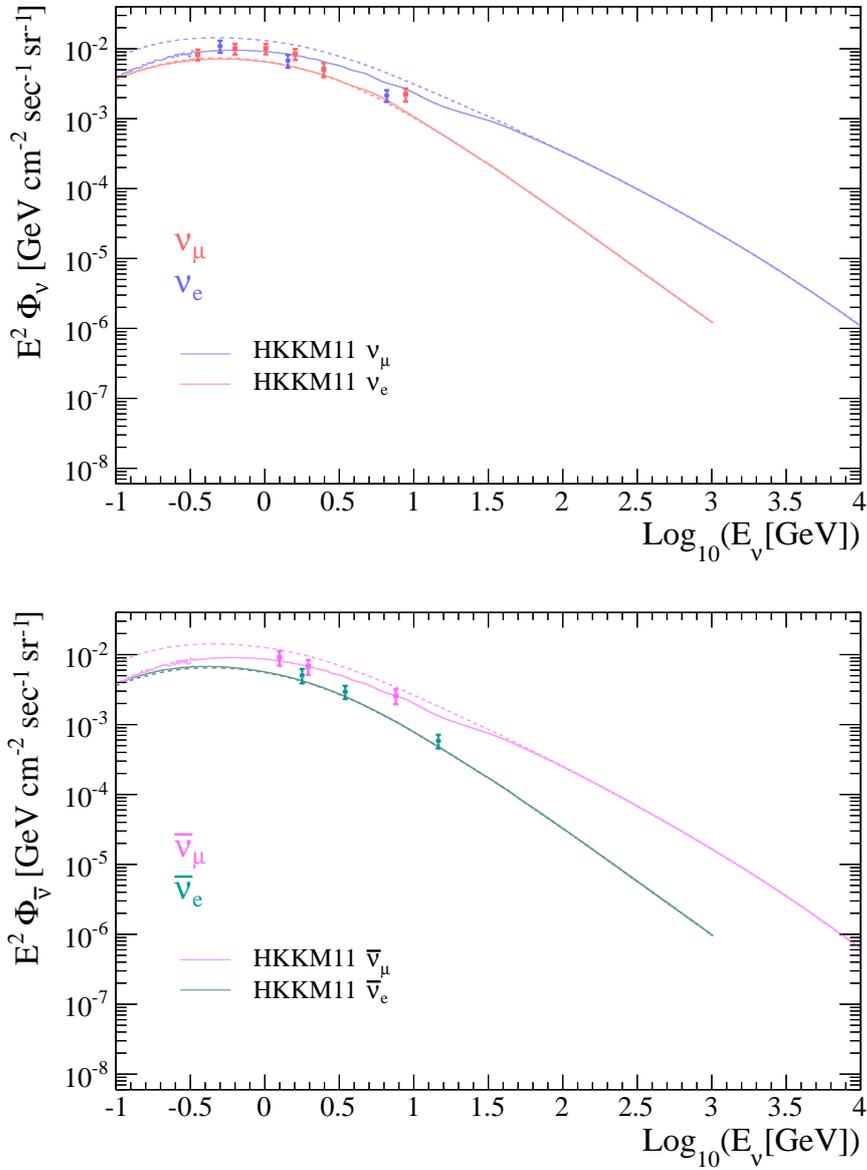
Figure 49: The measured ν (top) and $\bar{\nu}$ (bottom) fluxes, compared with the oscillated (solid line) and unoscillated (dashed line) HKKM predictions.
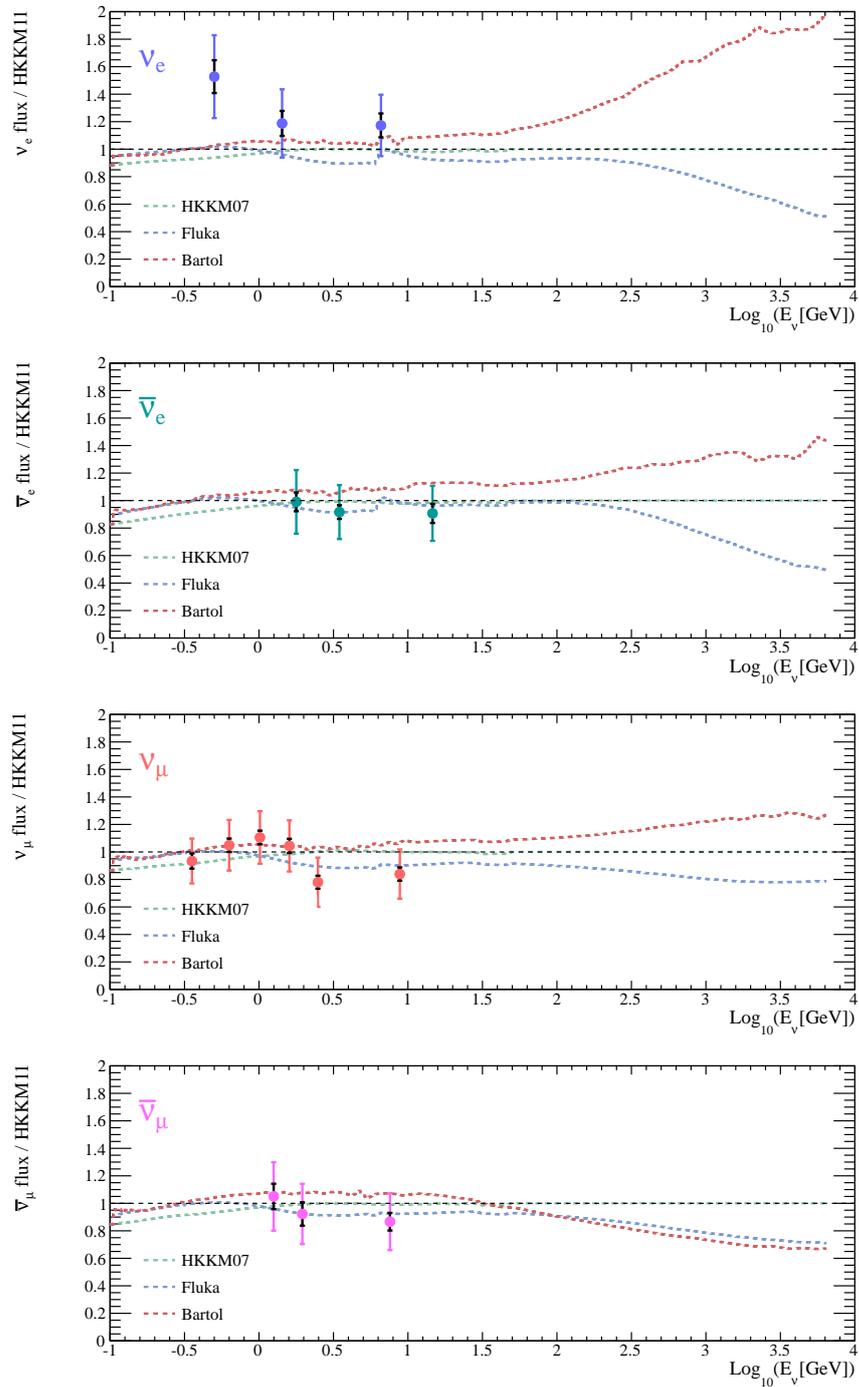
Figure 50: The measured flux ratios between the data and HKKM11 model, for each neutrino type, with other models overlaid. Colored error bars represent statistical and systematic uncertainties combined, while black error bars represent statistical uncertainties only.

MEASUREMENT OF THE AZIMUTHAL SPECTRA OF THE NEUTRINO FLUX

6.1 INTRODUCTION

Neutrino experiments have often focussed on the zenith angle distributions of the neutrino flux, as neutrino path-length is a strong function of the zenith angle, and the oscillation effects may be investigated. However, the azimuthal distributions also contain some interesting physics. The geomagnetic rigidity cut-off of primary cosmic rays at the Earth's surface is an anisotropic function, as discussed in Sec. 2.2.2, where Fig. 7 showed the rigidity cut-off at the SK site in detail as a function of zenith and azimuthal angles. The strongest effect of this anisotropy is the well-known suppression of west-going primary cosmic rays (illustrated clearly in Fig. 6), causing an east-west anisotropy in the secondary cosmic ray flux. This effect is of course expected to appear in the neutrino flux, where it has been previously discovered for ν_e, but indicated only with low significance ($< 2\,\sigma$) for ν_μ, in a previous Super-Kamiokande measurement [135].

Recent atmospheric flux models have also progressed from one-dimensional to three-dimensional calculations, including modeling the bending of cosmic ray secondary particles in the geomagnetic field, which has led to significant modifications of the predictions of the azimuthal distributions [19]. While the dominant effect remains the primary west-going suppression, the asymmetry is enhanced by an east-west dipole effect caused by muon bending, with a strength that depends strongly on the energy and zenith angle. Measuring in detail the azimuthal distributions can thus be used as a cross-check to the flux simulations, in particular to test the accuracy of their implementations of geomagnetic effects; due to the increased statistics since the previous measurement, we may also have the power to discover the expected azimuthal anisotropy in the ν_μ flux.

6.2 EVENT CLASSIFICATION AND ANALYSIS METHOD

6.2.1 *Selection and binning*

The azimuthal analysis uses only the FC sample, as described in Sec. 4.2, and selects all FC e-like and μ-like single-ring events. These samples have a very high neutrino flavour purity, estimated as 94% ν_e for the e-like sample, and 98% ν_μ for the μ-like sample.

Each neutrino event is binned by the azimuthal angle ϕ, zenith angle θ, and reconstructed energy of the produced lepton E_{rec}. The azimuthal angle ϕ is defined as the clockwise angle between true south and the lepton's reconstructed forward-going direction, and is binned into 12 evenly-sized azimuthal bins from 0→360 degrees. The zenith angle θ is binned as 5 evenly-sized $\cos(zenith)$ bins from -1 to +1, where +1 represents an up-going lepton. The reconstructed energy E_{rec} is binned into 4 energy bins starting at 0.1 (0.2 for μ-like), 0.4, 1.33, and 3.0 GeV, where the last bin is unbounded from above. The energy resolution of the lepton per event, even in the lowest energy bin, is estimated as better than 7% (3%) for e-like (μ-like) events. This binning was chosen following the MC prediction, in order to show zenith and energy regions in which the predicted east-west effects have various different strengths, while keeping statistics high in each bin.

All data from the SK-I to SK-IV periods are summed together, for a total of 13,061 (e-like) and 12,711 (μ-like) events. The MC events are generated as described in Sec. 4.5, and binned in a similar way using their reconstructed variables.

6.2.2 *Asymmetry parameter*

The azimuthal distributions will be examined by plotting them for each energy bin while summing over zenith angle, and for each zenith bin while summing over energy. To quantify the east-west dipole asymmetry in each plot, we define the parameter

$$A = \frac{n_{east} - n_{west}}{n_{east} + n_{west}} \tag{63}$$

where n_{east} (n_{west}) represents the number of events with azimuth angle between 0→180 (180→360) degrees.

To calculate the significance of a nonzero east-west effect, we test a reduced sample containing the middle two energy bins and middle three zenith bins, which is to say all events in the 0.4→3.0 GeV energy range with $|\cos(\theta)| < 0.6$. These criteria were optimized by MC to select the event sample giving the largest predicted value of $A/\Delta A$, which is defined as the significance in units of σ.

6.2.3 *Zenith dependency of asymmetry shape*

The primary shape of the the rigidity cutoff effect is approximated as an east-west dipole shape, as can be clearly seen in Fig. 7 on page 22. However, away from the horizontal, it can be seen that the rigidity cutoff takes on a slightly more complicated shape. While the A parameter measures the strength of the east-west dipole asymmetry depending on energy and zenith, we also wish to test if the shape of

this asymmetry is affected by this zenith dependency of the rigidity cutoff shape.

In principle we could measure the relative flux of the neutrino as a complicated function of zenith and azimuth and compare with the rigidity cutoff map, but in practice the statistics are not high enough to accurately measure such differences. We can however define a coarser parameter, to attempt to measure some zenith dependency of the neutrino flux's azimuthal distribution shape. According to the MC prediction, it should be possible to measure these second-order effects by the phase-shift parameter B in the fitting

$$k_1 \sin(\phi + B) + k_2 \tag{64}$$

for each zenith plot (with k_1, k_2 free parameters). This parameter essentially fits the azimuthal alignment of the asymmetry dipole depending on the zenith angle, and is shifted away from zero for a non east-west alignment of the dipole. By toy MC, we see on average a 2.0 σ significance of a zenith-dependent B parameter, when combining both e-like and μ-like samples.

The actual fitting method of the data to the MC prediction uses the B parameter as a function of the five zenith bins (shown later in Fig. 60 on page 132), plotted separately for e-like and μ-like events. We define our hypothesis H for the bin contents as the shape predicted by MC, but also allow a free-floating overall normalization α_1 and y-axis scaling α_2 of the MC shape separately for e-like and μ-like events, such that we have four free parameters $H = H(\alpha_1^e, \alpha_2^e, \alpha_1^\mu, \alpha_2^\mu)$. We then define, for data counts N, the statistic

$$\Delta\chi^2 \equiv 2\ln\frac{L_M(N|H)}{L_M\left(N|H(\alpha_2^e = 0, \alpha_2^\mu = 0)\right)} \tag{65}$$

where L_M is a Gaussian maximum-likelihood estimator, which finds the best-fit α parameters.[1] In the denominator however, we restrict the α_2 parameters to be zero, which is equivalent to fitting a constant function for the e-like and μ-like plots – i.e. no zenith dependency of B. Since the difference in parameter space between the hypotheses in the numerator and denominator is two, by Wilks' theorem (as explained in Appendix B) the statistic $\Delta\chi^2$ should be distributed as a $\chi^2(k = 2)$ distribution (where k is the number of degrees of freedom). We may thus extract the significance of rejecting the no-zenith-dependency model as $P\left(\chi^2(k = 2) \geq \Delta\chi^2\right)$.

---

1 We use the notation $\Delta\chi^2$, as for Gaussian L_M this test is equivalent to the difference in the χ^2 statistic of the two hypotheses.
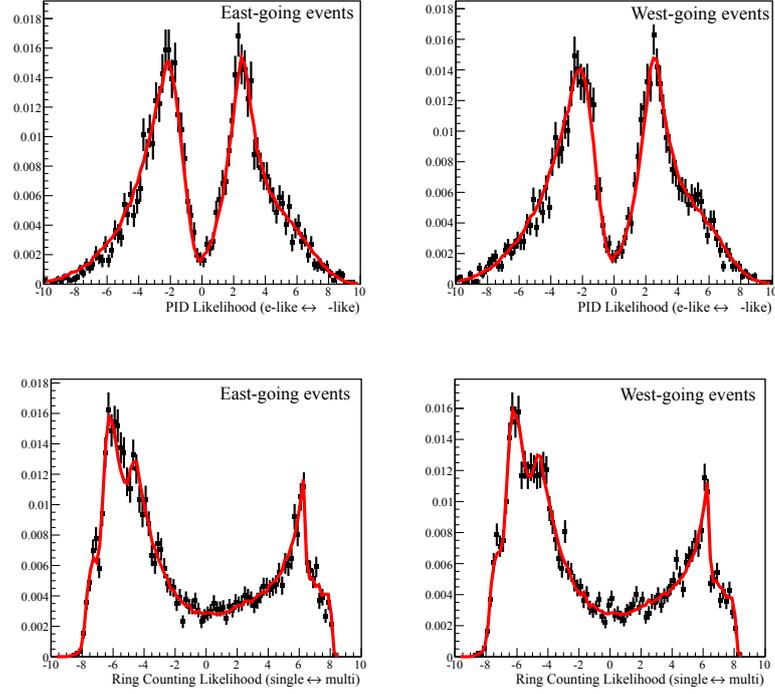
Figure 51: The ring counting and PID likelihood variables for the SK IV Sub-GeV data, separated into east and west-going events, shown with best-fit spline function based on the MC distribution.

6.3   SYSTEMATIC UNCERTAINTIES

6.3.1   *Azimuthally-dependent detector systematics*

The SK detector is considered to be highly azimuthally symmetric, and no azimuthally dependent systematic reconstruction errors are implemented in the SK systematic error database (as was described in Sec. 5.4 and summarized in Appendix A). The possible existence of such errors was nonetheless investigated.[2]

6.3.1.1   *PID and ring-counting*

Possible particle identification and ring-counting biases were checked using the distribution of their likelihood variables, for each detector period. The nominal likelihood shapes from the MC events were fitted to two sets of data, one for east-going events and one for west-going events, for sub and multi-GeV samples separately. Figure 51 shows an example of such fittings for the SK-IV sub-GeV sample, where the

---

2 A possible cause of such errors could be, for example, a residual magnetic field inside of the SK tank. While the geomagnetic field is in principle compensated by Helmholtz coils, a previous measurement of the remaining field throughout the tank showed that while the mean field in the vertical direction was less than $0.1\,\mathrm{mG}$ a small horizontal mean field at the ~$30\,\mathrm{mG}$ level remained, which was expected to effect PMT collection efficiency at the 1~2% level [79].
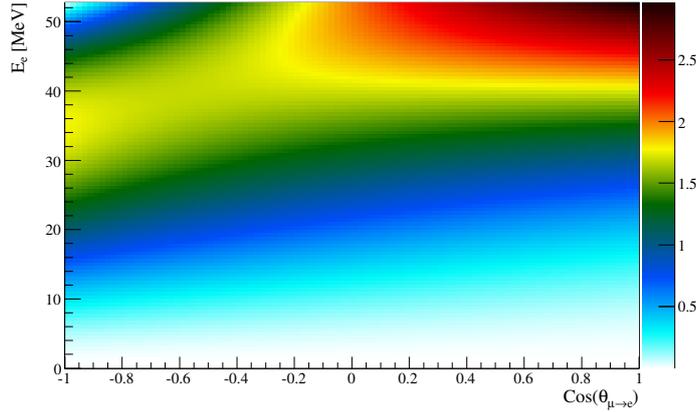
Figure 52: The predicted decay rates for the Michel decay $\mu^- \to e^- + \bar{\nu}_e + \nu_\mu$ ($\mu^+ \to e^+ + \bar{\nu}_\mu + \nu_e$). Here E_e is the electron (positron) energy, and $\theta_{\mu \to e}$ is the inner angle between the muon *spin* and electron (positron) *momentum*. Column normalized.

shape was fitted allowing a shift or a re-scaling along the *x*-axis. The fitted shapes of the distributions were consistent between the east and west-going events, and the error in either of these parameters was considered negligible.

6.3.1.2 *Reconstructed energy*

The variance of reconstructed momentum scale by azimuth was checked using electrons (and positrons) coming from the decay of cosmic-ray muons which come to a complete stop in the detector. The energy spectrum of such electrons is the well-known Michel spectrum. Such a sample can be selected by requiring a fitted muon track that comes to rest in the detector, and a single contained electron-like event that follows, in addition to strict cuts on the decay vertex position, decay time, and muon and electron fitter goodness.

However, there are several factors relating to the muon polarization that must be corrected for. Figure 52 shows the muon decay rate as a function of the the angle between the muon spin and the electron momentum, and the electron's energy. It can be seen that in the forward direction of polarization, the electron tends to be given more kinetic energy. Thus, if the muon flux at SK has an azimuthally-dependent polarization, we should expect a real energy bias in the Michel electrons that must be corrected for. In fact, even an azimuthally-independent average polarization will cause an azimuthal bias, as the azimuthal decay distribution will be weighted by the muon flux intensity, which differs depending on the azimuth angle due to the shape of the mountain above SK.

At least some muon polarization should be expected considering the Lorentz boost between the parent frame of the atmospheric π
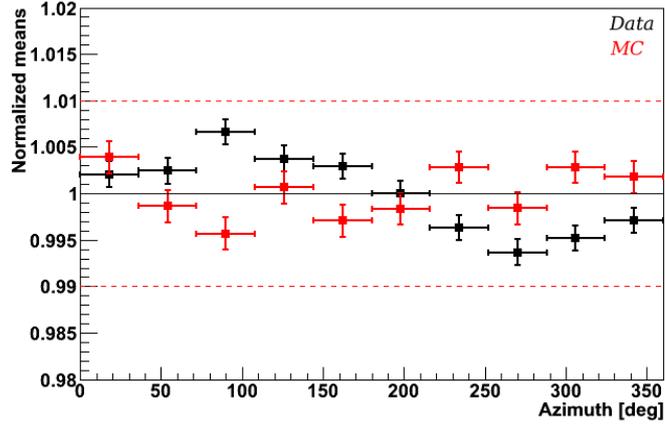
Figure 53: The mean energy of decay-electrons depending on the azimuth angle, with the average energy normalized to one.

or K and the lab frame. Previous measurements by the Kamiokande detector [136] (which take into account capture and depolarization of μ^- in water) suggest an average forward muon polarization of 0.26, which is implemented into the stopping-muon MC. In order to reduce this polarization effect, we implement a cut requiring the stopping muon direction to be within 20° of down-going, which is the tightest cut possible while keeping reasonable statistics.

The final result of the azimuthal energy scale check is shown in Fig. 53 for MC and data. While the MC shows no clear systematic bias, the data shows a small and roughly sinusoidal azimuthal bias of the order 0.6%. Although this may be plausibly due to a poor modeling of the muon polarization, it may be due to a genuine bias in the detector. An azimuthal energy calibration systematic is therefore implemented as a sinusoidal effect, with a magnitude of 0.6% for a 1 σ value of the error strength. The final effect of this systematic on our data sample is shown later in Tab. 17.

6.3.2 *Second-order systematics due to azimuthal flux asymmetries*

Considering the neutrino flux at SK, the energy spectrum, average path length, and $\nu/\bar{\nu}$ ratio, are expected to change slightly depending on the azimuth angle [68] due to the geomagnetic effects.[3] Any detector systematic that depends on these factors will thus be azimuthally dependent, in second-order, and must be considered.

To give a clear example of what we mean by a "second-order" dependence, consider for example the MC true energy spectra of μ-like events shown in Fig. 54. We see that east-going sub-GeV events have on average a slightly lower energy than west-going events; since some

---

3 These changes are expected to be too small to be directly measurable, which is why this study considers only on the flux normalization as a function of azimuth.
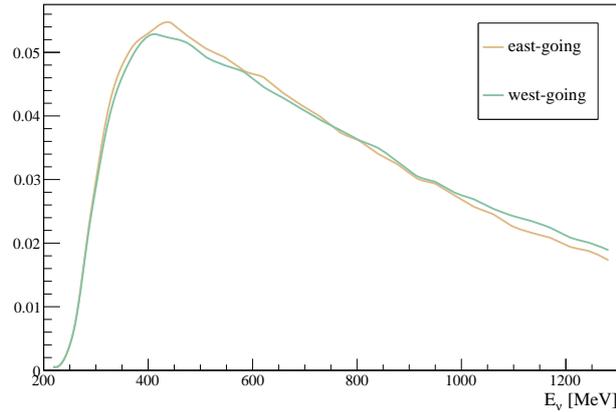
Figure 54: The energy spectra for the μ-like single-ring sub-GeV events, estimated by MC, separated into east-going and west-going events. Plots are normalized to unity.

detector reconstruction uncertainties are larger at those lower energies, a larger systematic error is applied to the east-going events.

Many systematics in the SK systematics database, described in Sec. 4.5.4 and Appendix A, are a function of neutrino energy. The systematics software was modified to process events as a function of azimuthal angle (in addition to zenith angle and neutrino energy), and we calculate their effects at 1σ in the azimuthal binning, including cross-section errors (15), detector reconstruction errors (53), neutrino oscillation model errors (6), and flux model errors (21). The detector-related errors are numerous as most of them are modeled separately for each SK period, however since the selection of single-ring events is simpler than the multiple sub-samples used in the previous energy spectrum analysis, comparatively less reconstruction-related errors are applicable.

6.3.3 *Implementation of the systematics*

For the parameter A in Eq. (63), any systematic shift in the total event number that applies equally to east-going and west-going events will cancel out, and give no contribution to the final error estimation. Conversely, any systematic that effects east-going and west-going events differently will give some contribution to the final error.

Similarly, when we plot the full azimuthal distributions, we are interested in testing their shape as a function of azimuth, rather than measuring the absolute normalization of the neutrino flux. Thus if we consider a systematic effect that applies equally to all azimuthal bins, it should not contribute to the systematic error estimation. For each systematic error, we define it's "normalization" component as the average effect on all azimuthal bins at, and it's "azimuthally-dependant"

component as the effect on each bin relative to the average effect, at the 1σ level. The normalization component is discarded, and the azimuthally-dependant components of all systematics are combined in quadrature to calculate the final systematic error in each azimuthal bin.[4]

The systematics with the strongest effects on the azimuthal analyses are listed in Tab. 17. Cross-section errors are still dominant, and there is some contribution from relative flux normalizations, even though these errors are only azimuthally dependent in second order (as described in the previous section). Some detector reconstruction uncertainties for SK-IV and SK-I are also noticeable, due to their high livetimes compared with other SK periods.

6.3.4   *Azimuthal angle reconstruction*

Finally, it is important to note that the dominant factor in mis-reconstructing the neutrino direction does not come from detector inadequacies, but from the fact that in CCQE interactions at lower energies the neutrino direction and the produced lepton direction are poorly correlated. This is not modeled as a systematic error, but the effects are generally accounted for by the high statistics of the SK MC, when used in forward-fitting procedures. The correlation for reconstructed single-ring events is plotted by energy in Fig. 55, which shows that for interactions at less than 400 MeV there is only a faint ability to even discriminate the neutrino's forward direction from its backwards one. Events in the range $400 < E_\nu < 1330$ are generally correlated at least within 90°, which allows a good separation between east and west-going events. Finally, multi-GeV events are very well correlated within $\sim 15°$.

6.4   RESULTS AND DISCUSSIONS

6.4.1   *Azimuthal distributions*

The azimuthal distributions, using the lepton's reconstructed energy and direction, are shown in Fig. 56 for data and MC events. The values of the dipole anisotropy parameter A corresponding to each plot are shown in Fig. 57.

6.4.1.1   *Monte Carlo distributions*

To understand the origin of these distributions from the true flux shape, we also show the same MC events rebinned by true neutrino direction in Fig. 60, and first discuss only the MC predictions.

---

4 For a meaningful comparison, the MC must also normalized to the data in this case; the required normalization change is less than 2%.

SYSTEMATIC ERROR	SIZE OF EFFECT [%]
Neut axial mass parameter	0.59
MC statistical error	0.49
CCQE cross-section ratio	0.46
Axial mass parameter in QE and single π	0.35
Flux relative normalization $<$ 1 GeV	0.34
Single meson cross-section	0.34
Flux relative normalization $>$ 1 GeV	0.25
Neutral / charged pion ratio	0.22
Coherent pion cross-section	0.17
Deep inelastic scattering Q^2 at low W	0.10
Relative normalization for sub/multi-GeV FC	0.10
Single pion $\nu/\bar{\nu}$ ratio	0.08
CCQE $\nu/\bar{\nu}$ ratio	0.08
Deep inelastic scattering model differences	0.08
Δm_{23} error (from T2K)	0.07
Azimuthal energy calibration (SK IV)	0.06
Fiducial volume (SK IV)	0.06
$\nu/\bar{\nu}$ ratio $1 < E_\nu < 10$ GeV	0.05
Overall energy calibration (SK IV)	0.05
Azimuthal energy calibration (SK I)	0.05
Others	0.88

Table 17: Each systematic error and the total effect of their azimuthal component, given as a percentage shift away from the central MC values for a 1σ shift in the error, when binned in the azimuthal analysis binning. "Others" represents the sum of 74 systematics with individual effects $< 0.05\%$.
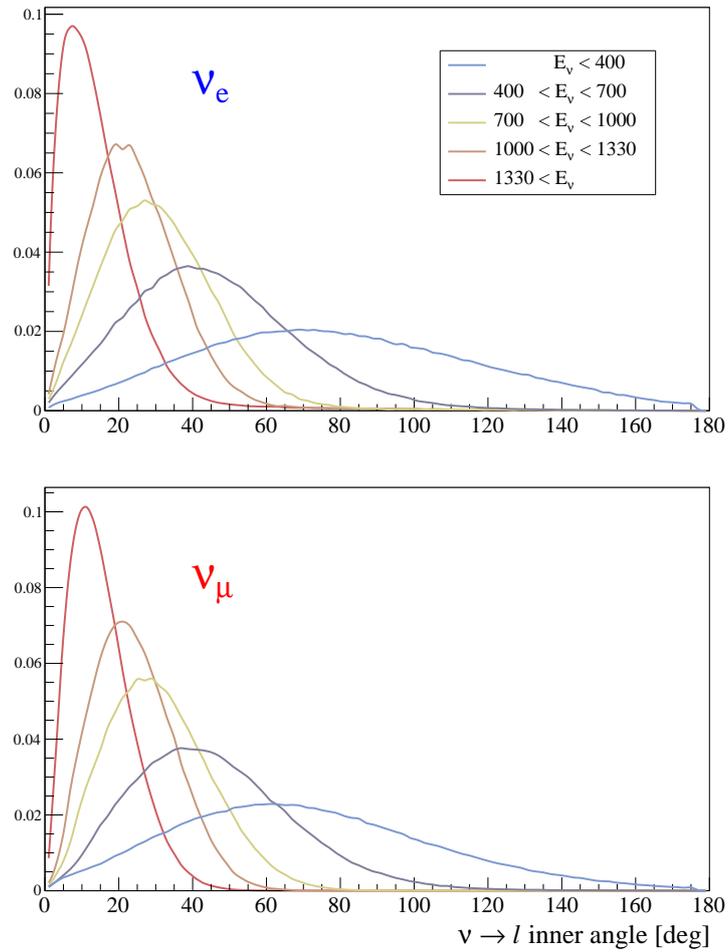
Figure 55: The neutrino → lepton true direction correlation for the single-ring events used in the azimuthal analysis, showing ν_e (top) and ν_μ (bottom). Neutrino energies E_ν are in MeV.

The MC truth plots show that at $E_\nu < 0.4\,\text{GeV}$ the true anisotropy is high, however the MC reconstructed plots show a diluted anisotropy, as was expected due to the poor correlation between the incoming neutrino and outgoing lepton directions. The anisotropy is strongest in the MC reconstructed plots at around $\sim 1\,\text{GeV}$, and it also for events around $\cos(\theta) = 0$.

The $\cos(\theta) > 0.2$ bins the MC truth plots show a more complicated shape, which is due to the rigidity cutoff having a suppressive effect in both east and west directions for up-going neutrinos. As this effect is strongest for sub-GeV events within $\sim 10°$ of up-going, it is also washed out in the MC reconstructed plots.

6.4.1.2 *Data distributions*

Considering now the observed data events, we see that the current data give excellent agreement with the MC expectations (based on the HKKM11 flux model). The χ^2 value (based on a weighted vs. unweighted events χ^2 test [137]) for the energy (zenith) distributions in Fig. 56 is 87.6 for 96 bins (106.6 for 120 bins), which for the MC prediction as the null hypothesis gives a p-value of 0.69 (0.79). The flux simulation is thus strongly consistent with the current data.

The final azimuthal distributions, using the reduced sample optimized for discovery of the asymmetry effects, are shown in Fig. 59. The final A parameters are found to be $A_e = 0.153 \pm 0.015(\text{stat}) \pm 0.004(\text{syst})$ for e-like events, and $A_\mu = 0.108 \pm 0.014(\text{stat}) \pm 0.004(\text{syst})$ for μ-like events. The east-west effect is thus seen at a significance level greater than 6σ (8σ) for the μ-like (e-like) samples. This is the first time it has been seen convincingly in the ν_μ flux.

6.4.2 *Zenith dependency of asymmetry shape*

Figure 60 shows the MC and data results for the B parameter, were the data matches well with the MC, and shows a 2.2 σ indication for the existence of a zenith dependence of the azimuthal asymmetry shape (using the analysis method as explained in Sec. 6.2.3, combining the significance from the e-like and μ-like samples together). This is the first indication that the geomagnetic effect produces an asymmetry which is more complicated than the "east-west" effect that is usually assumed.

In future, detectors such as Hyper-Kamiokande should be able to perform an improved test of the B parameter. By assuming similar systematic errors, but increasing statistics by a factor of 10 compared to SK (corresponding to a few years of observation by HK), we find the estimated sensitivity to observe a non-zero B parameter as 5.9 σ.
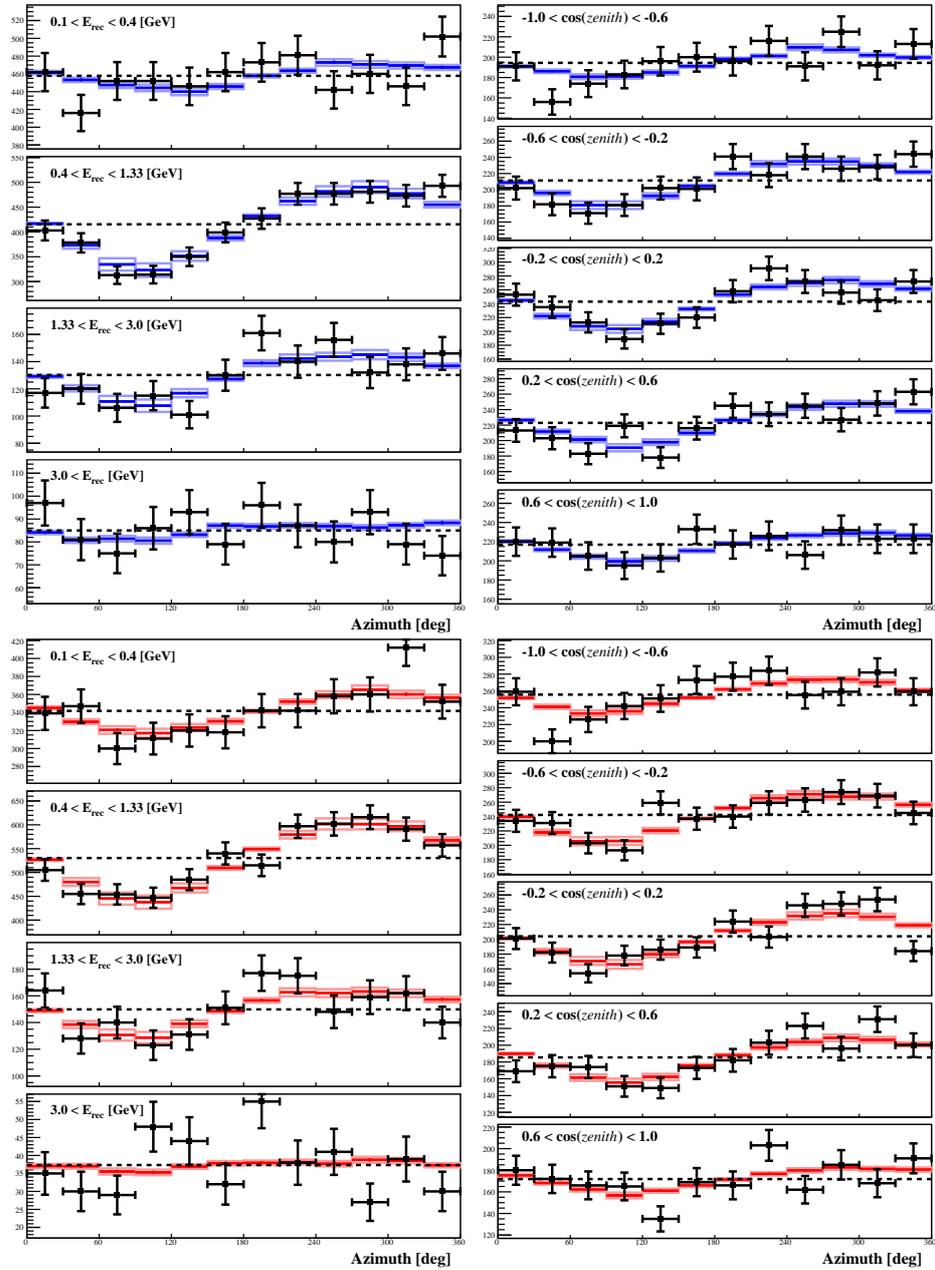
Figure 56: Single-ring e-like (top) and μ-like (bottom) events for the SK-I to SK-IV data (error-bars) and MC (error-boxes), by reconstructed azimuthal angle. The left plots show reconstructed energy ranges summing over all zenith angles θ, and right plots show zenith angle ranges summing over all energies E_{rec}. The plots are zero-suppressed to show the data/MC shape comparison clearly.
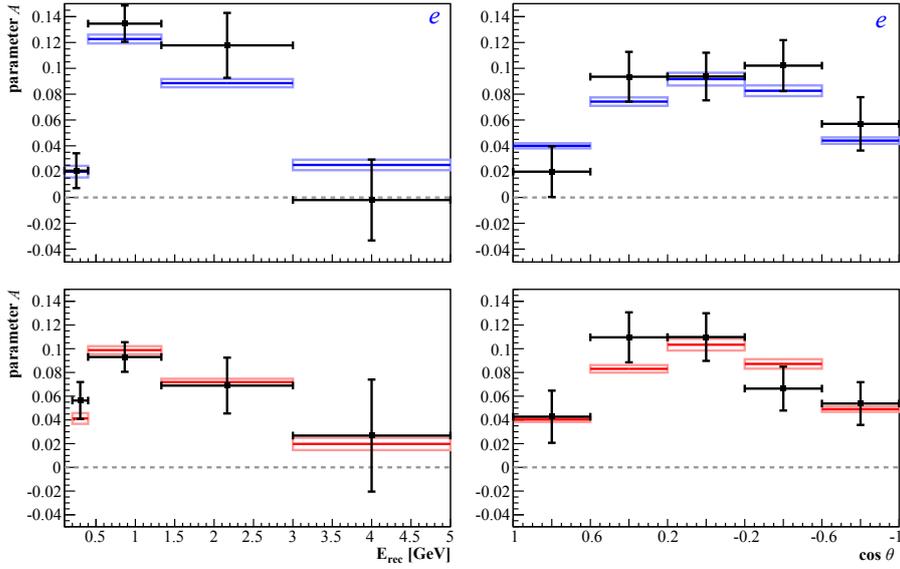
Figure 57: The parameter A depending on zenith angle or energy, for the e-like (top) and μ-like (bottom) events, using the SK-I to SK-IV data (error bars) and MC (error-boxes). The highest energy bin is unbounded from above, but is plotted only to 5 GeV.
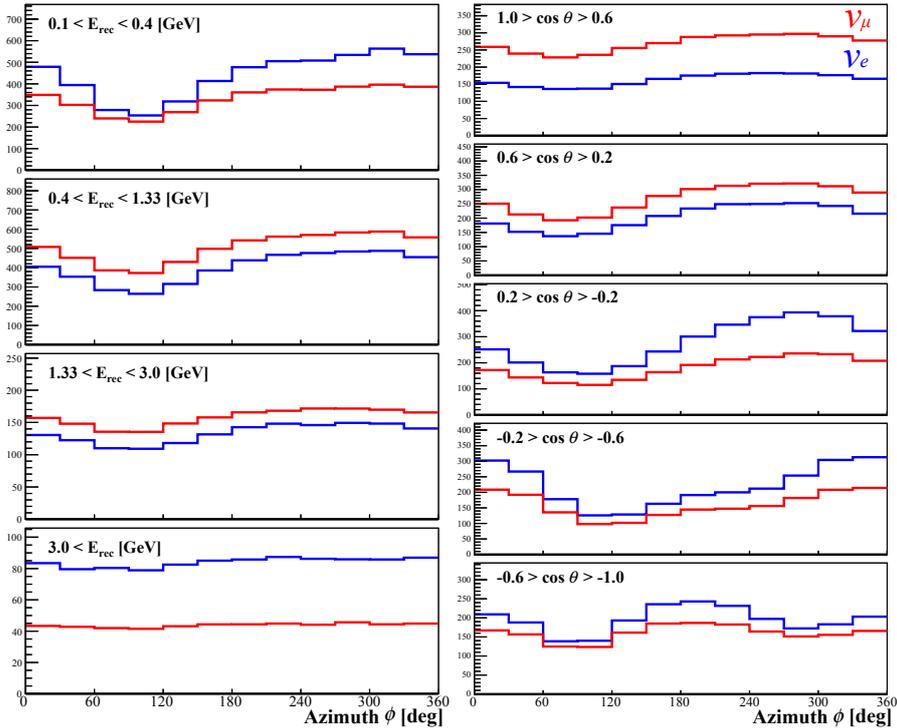
Figure 58: Reconstructed e-like (blue) and μ-like (red) MC events, as in Fig. 56, but re-binned using the MC true neutrino direction (for both zenith and azimuth angles) instead of the reconstructed lepton direction.
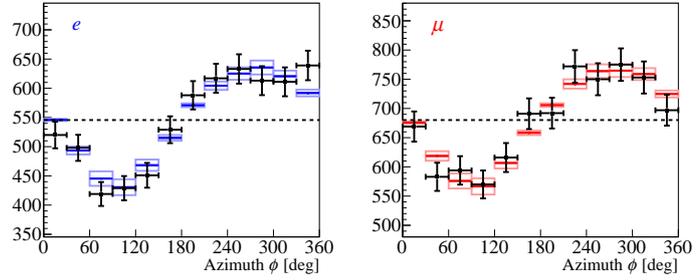
Figure 59: e-like (left) and μ-like (right) events for the SK I-IV data (error-bars) and MC (error-boxes), optimized for calculation of the final A parameters, selected by $0.4 < E_{rec} < 3.0$ GeV and $|\cos(zenith)| < 0.6$.
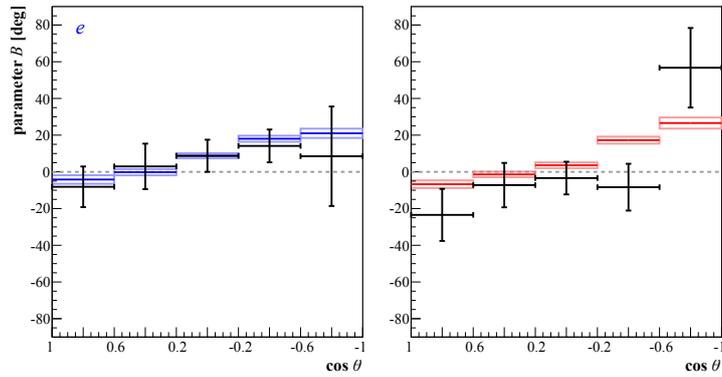
Figure 60: The parameter B depending on zenith angle, for the e-like (left) and μ-like (right) single-ring events using the SK-I to SK-IV data (error bars) and MC (error-boxes).

6.4.3 *Flux model comparisons*

Azimuthal flux distributions were also published by the Bartol group in 2003 [21]. Although this is older than the current HKKM group's model, the major relevant improvements (a three dimensional calculation with bending of secondary cosmic rays in the atmosphere) were included. The Bartol data contains 18 azimuthal bins, but only a single zenith bin from $-0.5 < \cos(\theta) < 0.5$. By reweighting the MC from the default HKKM11 azimuthal spectra to the Bartol azimuthal spectra (for events covered by the above range only), we created a new set of plots similar to Fig. 56, and obtained a χ^2 statistic for the Bartol model. The largest differences (up to 25%) in the azimuthal spectra of the flux models occur at $< 400\,\text{MeV}$, but due to the poor neutrino-lepton directional correlation, the observable differences in that range are small. The HKKM11 model has a χ^2 value lower by 1.0 units, but considering the number of bins such a difference seems insignificant, and we cannot draw any strong conclusions on flux model preference from the azimuthal distributions.

7

MEASUREMENT OF THE CORRELATION BETWEEN THE SOLAR CYCLE AND THE NEUTRINO FLUX

7.1 INTRODUCTION

The solar magnetic activity cycle is an oscillatory change in the solar density and structure, which causes a change in the level of solar plasma emissions, with an average period of approximately 11 years. The cosmic ray flux at earth is well known to be dependent on the solar activity [138]. This is because the plasma flux, or "solar wind", acts as a barrier by scattering cosmic rays entering the solar system and approaching the Earth; therefore during solar activity maxima the cosmic ray flux is reduced. Consequently the atmospheric neutrino flux is predicted to reduce – although this has not yet been measured.[1]

Historically, the solar activity was measured by its correlation with the appearance of sunspots, which are regions of intense surface magnetic activity causing a drop in temperature in that region, often easily visible from Earth. Since 1948, the use of neutron monitors (NMs) provides a method to accurately and constantly track the neutron flux at the Earth's surface resulting from cosmic ray impacts [139]. As the primary cosmic ray flux is measured by devices mounted on balloons or spacecraft (which have limited flight times), it is not itself constantly measured, but the NM counts are generally believed to be well-correlated with the primary cosmic ray activity. Some corrections may however be necessary depending on the measurement site; firstly the relative atmospheric pressure, which alters the mean free path of particles in the atmosphere, must be corrected for. The down-going rigidity cutoff (an effect discussed in Sec. 6.1) at the sites must also be considered, as the plasma wind affects the low-energy primary flux more strongly, and areas with a low rigidity cutoff have a relatively higher flux at lower energies.

We thus plan to test for a correlation of the atmospheric neutrino flux with the solar cycle, by searching for a correlation between the neutrino flux at SK and the neutron detection rates at various NMs operated by other institutes. This method is able to test on very short timescales of $O(1\,\mathrm{h})$, as both neutrino and neutron observations take place on Earth, and the propagation of the solar wind within the solar system (with speed of the order $100\,\mathrm{km\,s^{-1}}$) need not be considered.

---

1 As was mentioned in Sec. 2.3.2, a modulation of the neutrino flux is also expected due to seasonal atmospheric variations, however at the SK site these variations are at the $\ll 1\%$ level [68] and assumed to be negligible in this analysis.
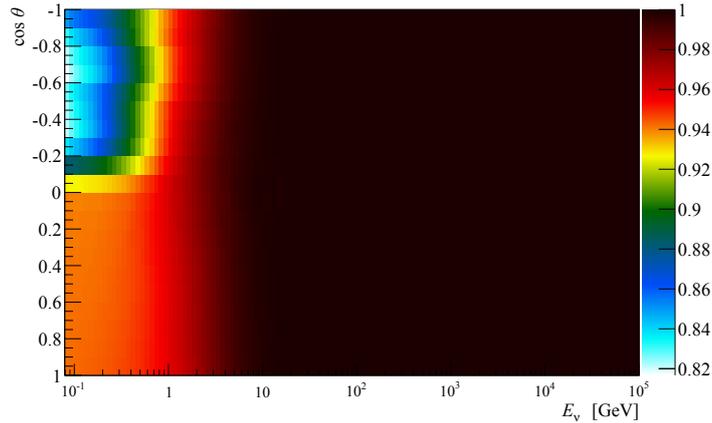
Figure 61: The reweighting function for the ν_e flux taken from the HKKM model, for the highest modeled solar activity corresponding to a Climax NM count of 3500 counts hr^{-1} ×0.01. Here θ is the zenith angle of arrival at SK.

We assume that at a given time the neighborhood of the Earth can be treated as a heat bath with constant temperature, and we may expect a good correlation between the neutrino flux at the SK site and the neutron flux monitored at the NMs in various locations around the Earth.

7.1.1 *HKKM flux model predictions*

Data was obtained from the HKKM flux group [74] that predicts the relative normalization change of the ν_e, $\bar{\nu}_e$, ν_μ, and $\bar{\nu}_\mu$ fluxes depending on the solar cycle. This data uses the count rate of a specific NM, the Climax NM [140], as the parameter corresponding to the degree of solar activity. Such a single parameter may be used as the model assumes that after corrections for the local air pressure at the monitor, all NMs are linearly correlated with the solar flux modulation – even though if two monitors are located at different rigidity cutoffs, the gradient of the correlation would be expected to be different. The HKKM model then calculates the effect at the SK site, extrapolating the effect of the solar modulations in all directions, where rigidity cutoff is a function of direction.

As an example, Fig. 61 shows the expected reweighting required for the ν_e flux at the SK site during the highest solar activity covered by the model. This corresponds to a Climax NM count of 3500 counts hr^{-1} ×0.01 (where the factor of 0.01 in the count rate is the conventional notation for the NM data), which was the average count rate over a period roughly corresponding to the 22nd solar maximum.[2] Such

---

2 Here "22nd" refers to the cycles measured since sunspot counting began on Earth; the 22nd cycle ran from 1986 to 1996.

reweighting functions are provided at intervals of 50 up to 4150 counts hr^{-1} $\times 0.01$, which is not the absolute maximum value recorded by the Climax NM, but rather a solar activity below which the change in ν flux is calculated to be negligible and the reweighting function approaches unity.

From Fig. 61, we see that only the neutrino flux below 1 GeV is strongly affected, so we may focus our analysis on the SK "sub-GeV" event samples as described in Sec. 4.2.2, which have reconstructed lepton energies $E_{rec} < 1330$ MeV. It is also seen that there is a strong zenith-dependence of the flux suppression. Although the model assumes no directional asymmetry in the solar wind effects, a directional asymmetry occurs at the SK site due to the fact that the rigidity cutoff causes a relatively higher low-energy neutrino flux around the Earth's geomagnetic polar regions compared to the equatorial regions. Since the solar wind has a stronger suppressive effect on low-energy particles, the biggest relative flux reductions are in the polar regions. After considering the geometric effect of translating from global coordinates to zenith coordinates at the SK site, the predicted difference in the suppression effect appears as an up-down asymmetry.

7.2 EVENT CLASSIFICATION AND ANALYSIS METHOD

7.2.1 *Super-K data selection*

We select the sub-GeV samples only for this study, since the multi-GeV samples show a negligible solar modulation effect. Table 18 shows all sub-GeV samples, from which we further select only the single-ring *e*-like and μ-like samples. The other sub-GeV samples have either low purity, or high systematic normalization errors; as we fit the ν_e and ν_μ data according to separate models, and the solar modulation effect is at most a ~10% effect, these samples are essentially not useful. Studies were done to check if using an energy sub-range of the sub-GeV data could increase the statistical power, however no strong benefits were seen.

The data we use in this analysis are thus simply categorised as either *e*-like or μ-like, up-going or down-going (based on the fitted lepton direction), and by SK period – resulting in a total of $2 \times 2 \times 4 = 16$ samples. For this analysis we use the very latest SK dataset, which covers up to April 2015; compared to the analyses in Sec. 5 and Sec. 6, this gives an additional six months of data.

7.2.2 *Neutron monitor data*

We search for NM data [141] from monitors that have been active and well-maintained throughout the entire SK experimental period. While the Climax NM does not fall into this category (having shut down in
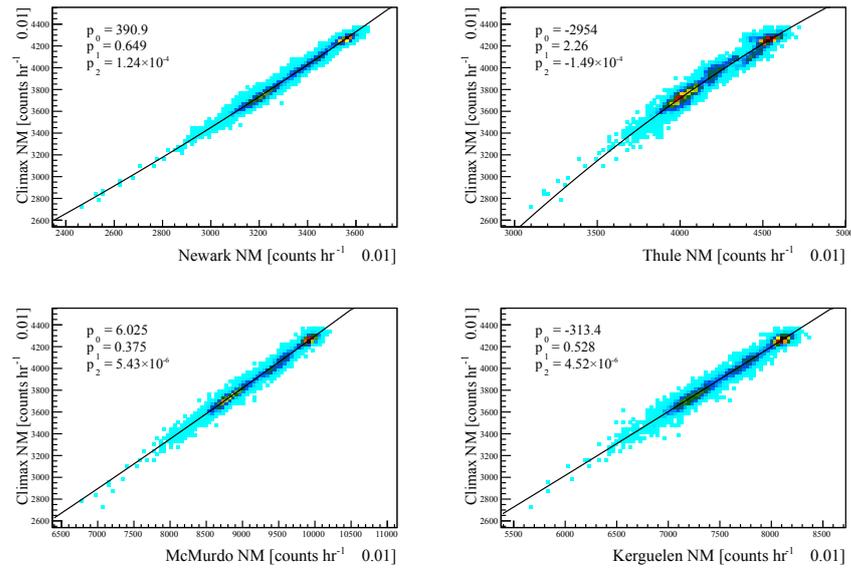
Figure 62: Fitted conversion functions between the pressure-corrected data from several NMs to the Climax neutron monitor. Each data point represents the neutron count recorded at each monitor during a given hour.

the 90s), we find four such NMs – the Thule, McMurdo, Kerguelen, and Newark monitors [129, 142]. These stations all monitor downgoing neutrons with relatively low rigidity cutoffs (0.3, 0.3, 1.14, and 2.4 GV respectively), and thus have good sensitivity to changes in the solar cycle, which as stated affects most strongly the low-energy cosmic ray flux. In particular, the Thule and McMurdo monitors in the north and south polar regions are very sensitive to the solar activity.

To obtain an "equivalent Climax NM count" for each of these NMs, we compare the counts of the monitors during times when they were both operational, as shown in Fig. 62. Although the correlation is almost as well fitted by a linear fit (as is assumed by the HKKM model), a second-order polynomial is used to account for a small non-linearity.

The χ^2-statistic per degrees-of-freedom, considering the statistical errors only, is however unsatisfactorily large on these plots, indicating some systematic differences in the NM monitors. To minimise systematic error, we define a "Climax NM parameter", which is the average of the four NM values after each was converted to the equivalent Climax NM count using the fitted polynomials. The systematic error on this final parameter at 1 σ is estimated, by taking the average RMS of the four counts, as 15.8 counts $hr^{-1} \times 0.01$.

The variance of this Climax NM parameter over the SK operating period is shown in Fig. 63, showing that with recent data included, almost two solar maxima are covered by SK data (despite unfortunately some downtime between SK-I and SK-II).
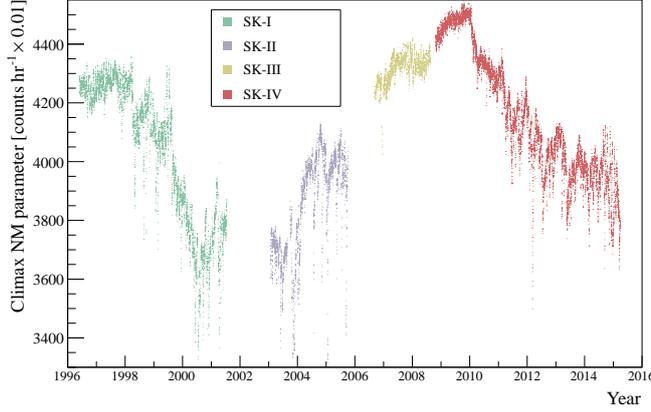
Figure 63: The values of the "Climax NM parameter" plotted over the operational history of the SK detector, where each point represents a neutrino event. Two solar maxima can be clearly seen, as can some areas where SK was not operational (such as the maintenance works between SK periods). Some sharp downward peaks can be seen, which will be analysed separately as described in Sec. 7.2.7.

7.2.3 *Parameterization of the effect*

Using the HKKM model, the SK MC events are reweighted based on their true neutrino energy and zenith angle, using reweighting functions such as those in Fig. 61. The fractional decrease in the number of events in each reconstructed sample, calculated using the SK-I→IV MCs weighted appropriately by livetime, is shown in Fig. 64. It can be seen that the suppression effect is not linear in the NM count parameter, and only has a significant effect compared to the relative normalization systematic errors of $\sim 3\%$ at NM count values below around 3700. Comparing with Fig. 63, We see that only the data obtained during the solar maxima at the end of SK-I and beginning of SK-II, and some the SK-IV data covering the most recent solar maxima, will be sensitive to the solar modulation. Of course, data in the low-solar activity periods is also essential, in confirming the expected event rates against which to compare any relative decrease.

The four functions in Fig. 64 give the prediction of the suppression effect in a more realistic way than a simple linear fitting would achieve for each of the four data samples, however it is possible that the effect is stronger or weaker than the prediction in reality. To test for this possibility when opening the data, we define a function $f_s(\alpha)$ for each sample s which takes its shape from the predicted functions, but allows them to be rescaled by the single continuous parameter α. We define $\alpha = 0$ to represent no solar activity and $\alpha = 1$ to represent the default prediction of the HKKM model, but higher or lower
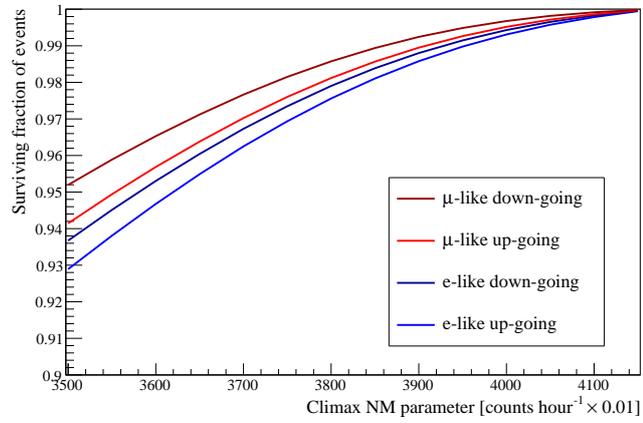
Figure 64: The effect of reweighting by solar activity on the SK-I→IV reconstructed data samples, calculated from the SK MC according to the HKKM model.
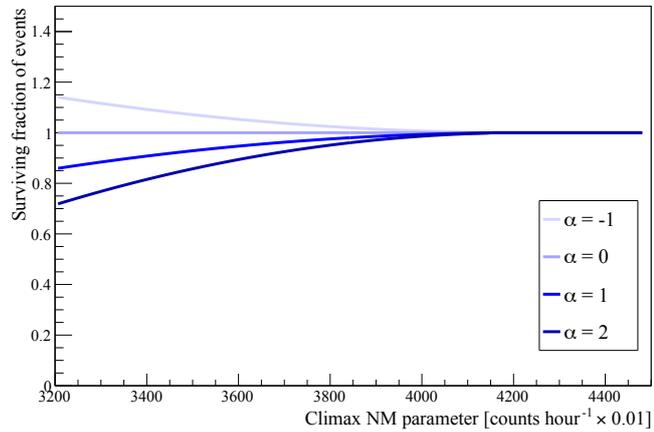
Figure 65: The function $f_s(\alpha)$ showing an example of the parameterization of the solar activity effect by a single parameter α, in the case where the sample s is the e-like up-going sample.

SAMPLE	EVENTS	PURITY [%]	ERROR [%]			
			SK-I	SK-II	SK-III	SK-IV
Single-ring e-like	10,892	95.0	2.80	2.82	3.00	2.35
Single-ring μ-like	10,763	98.2	2.53	2.87	3.13	2.48
Single-ring π^0-like	545	45.5	2.2	2.5	2.2	2.1
Multi-ring μ-like	823	93.6	5.2	8.9	7.5	4.2
PC	374	99.1	4.6	6.3	6.6	5.8

Table 18: The SK-I→IV data, counting sub-GeV events only. The purity is defined as the fraction of correct flavour interactions, estimated by MC (for the single-ring π^0-like sample, the "correct flavour" is assumed to be ν_e). The error represents the relative normalization error depending on SK period, calculated by combining all applicable systematic errors in quadrature, as described in Sec. 7.2.4.

values are also possible. As an example, functions for the SK-I e-like sample as a function of α are shown in Fig. 65. We thus have a family of functions that can simultaneously be fitted to the 4 data samples in each of the 4 SK periods, in order to measure the overall strength of the solar modulation effect by the single parameter α.

7.2.4 *Systematic errors*

As the solar modulation effect is a relative normalization shift as a function of the solar activity, any systematic error that affects the overall normalization of all SK data may be ignored. However, since the datasets of the SK-I to SK-IV periods are combined, we must consider errors that are dependent on detector period, and may cause a systematic shift in the relative normalization of each period. Such errors can arise from detector changes such as the differences in the PMT properties and distribution in the tank, or the replacement of the electronics. Table 18 shows the overall calculated systematic errors for all samples by detector period (although only single ring e-like and μ-like are used in the analysis).

The breakdown of the individual effect of each systematic is shown in Table 19. The dominant detector-dependent errors (>1% level) are the fiducial volume cut and the ring separation uncertainty; for the e-like sample the single ring π^0 fitter uncertainty also contributes, and for the μ-like sample the decay electron tagging uncertainty contributes (although this error is much reduced with the SK-IV electronics). While these errors are considered as completely independent for each detector period, in reality there may be some correlation for certain errors, which would mean that the error in the relative normalization would be somewhat conservative.

ERROR	EFFECT ON e-LIKE [%]				EFFECT ON μ-LIKE [%]			
	SK-I	SK-II	SK-III	SK-IV	SK-I	SK-II	SK-III	SK-IV
FC reduction	0.20	0.20	0.80	0.30	0.2	0.20	0.80	0.30
FC / PC separation	-	-	-	-	0.01	<0.01	0.01	< 0.01
Non-ν_e background	0.50	0.20	0.10	0.10	-	-	-	-
Non-ν_μ background	-	-	-	-	0.21	0.11	0.11	0.12
Ring separation	1.49	1.47	1.88	0.51	0.70	1.33	1.56	0.84
Single ring PID	0.23	0.66	0.26	0.28	0.18	0.50	0.19	0.22
Energy calibration	0.12	0.18	0.30	0.26	0.40	0.61	0.99	0.85
Up / down energy calibration	< 0.01	< 0.01	0.01	< 0.01	0.03	0.03	0.06	0.01
Azimuthal energy calibration	0.01	0.01	0.01	0.01	0.03	0.01	0.02	0.02
Single ring π^0 fitter	1.12	1.12	0.83	1.00	-	-	-	-
Decay electron tagging	< 0.01	< 0.01	0.01	< 0.01	1.29	1.35	1.32	0.78
Fiducial volume cut	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00

Table 19: Breakdown of SK period-dependent systematic error effects in the single ring e-like and μ-like samples.
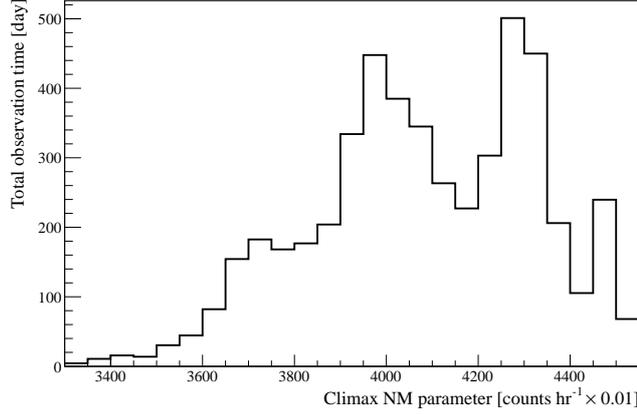
Figure 66: Total observation time t_i^{obs} at a particular NM count bin i (SK-I→IV inclusive).

As these errors are not dependant at all on the Climax NM parameter, we simply combine them in quadrature, and assign a total relative normalization error for each SK period, for the e-like and μ-like samples separately. There are thus a total of 8 systematic errors.

7.2.5 *Fitting method*

The histogram in Fig. 66 shows the total observation time in days t_i^{obs} for SK I-IV inclusive, binned by the Climax NM parameter averaged over 1-hour periods (with bins $i = 1 \rightarrow 25$). The minimum value on the Climax NM parameter in these plots is 3300 counts $\mathrm{hr}^{-1} \times 0.01$; there is a small amount of observation time below this value, but these data will be treated separately, as discussed later in Sec. 146.

In the case of no solar modulation effect, we can simply scale this histogram by the expected daily event rate for each sample r_s taken from SK MC, to calculate a prediction for data binned in the same way. In the case of a solar modulation effect, we must also apply the reweighting functions $f_s(\alpha)$ (as shown in Fig. 65). Explicitly, we define our hypothesis $H_{s,i}$ as the HKKM solar modulation prediction, given for each sample s in each bin i, by

$$H_{s,i}(\alpha) = t_i^{\mathrm{obs}} \times r_s \times f_{s,i}(\alpha) \qquad (66)$$

where $f_{s,i}(\alpha)$ is simply $f_s(\alpha)$ evaluated at the center of bin i. For example, Fig. 67 shows the number expected of events for the e-like up-going sample, for both the $H_{s,i}(\alpha = 0)$ and $H_{s,i}(\alpha = 1)$ hypotheses.[3]

---

3 On this plot, we should consider the uncertainty in the NM parameter, which may cause events to be binned incorrectly. This may affect the fitting due to the limited number of events at high solar activity. Even though such migration should be small
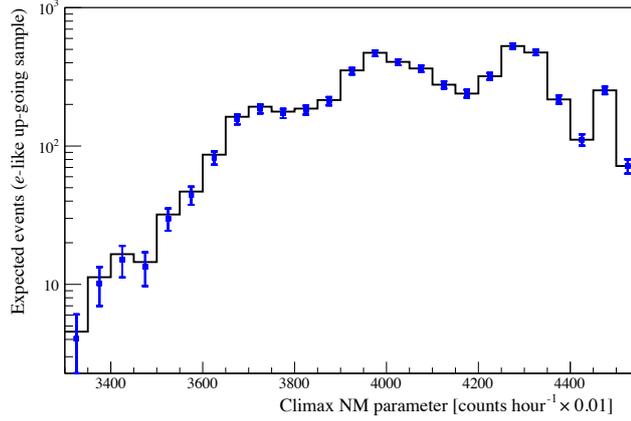
Figure 67: The predictions for the e-like up-going sample (one of four samples) in the case of no solar activity correlation ($H_{s,i}(\alpha = 0)$, black line), and of a correlation as predicted by the HKKM model ($H_{s,i}(\alpha = 1)$, blue points with statistical error).

Our 8 systematic errors (as discussed in Sec. 7.2.4) are taken into account by varying our hypothesis as follows. We define ϵ as the vector of systematic "error pulls", which holds the strength of each systematic in units of σ, and define a new hypothesis as

$$H'_{s,i}(\alpha, \epsilon, k_s) = H_{s,i}(\alpha) \times S_{s,i}(\epsilon) \times k_s \qquad (67)$$

where $S_{s,i}$ is the systematic modification, which has a strength depending on ϵ and a shape depending on the relative fraction of each SK period in each bin i, and k_s represents four free-floating overall normalization parameters, allowing a change in the the total ν_e and ν_μ flux normalizations (consistently across the SK-I to SK-IV periods).

The comparison between the two hypotheses $H'_{s,i}(\alpha)$ and $H'_{s,i}(\alpha = 0)$ and the data is mathematically performed as follows. Considering Fig. 65, we expect data counts in many bins low enough that the Gaussian approximation for statistical errors will be poor. In such a case we should consider directly the Poisson distribution for expected events, and use the likelihood ratio method to compare our MC and data. A brief mathematical description of this method is given in Ap-

---

since the error on the NM parameter (estimated as $15.8\,\mathrm{counts\,hr^{-1}} \times 0.01$ in Sec. 137) is much less than the bin size, it was estimated by unsmearing the distribution with a Gaussian function. This showed that any bias is negligible in the sensitive region, and should not contribute to the final result. In the areas where the observation time is quickly-varying as a function of the NM parameter, there may be some migration between bins at the $\sim 10\%$ level. However from e.g. Fig. 67, we can see that this happens only in the low solar activity region, and would have no effect on the fitted α-function which is flat in that region.

pendix B, which for data counts N_i leads us to consider the likelihood ratio

$$\Lambda \equiv 2 \ln \frac{L'_M \left(N_i | H'_{s,i}(\alpha, \epsilon, k_s) \right)}{L'_M \left(N_i | H'_{s,i}(\alpha = 0, \epsilon, k_s) \right)} \tag{68}$$

where L'_M is defined as

$$L'_M \equiv L_M + \epsilon^2 \tag{69}$$

where L_M is itself the maximum likelihood estimator described in Appendix B, which finds the best-fit for the parameters α, ϵ, and k_s of H' (the extra penalty term ϵ^2 is included here to constrain the systematic pulls, to avoid unrealistically high systematic effects). Importantly, in the denominator of Eq. (68), α is restricted to be zero, which means there is one less free parameter than in the numerator; by Wilkes' theorem, we can then expect that Λ is distributed as a $\chi^2(k = 1)$ distribution[4] and the significance to reject the no-solar-correlation model $H_{s,i}(\alpha = 0)$ can be taken as $\sqrt{\Lambda}$.[5]

7.2.6   Toy MC

Toy datasets were generated for each of the 16 samples, for both $\alpha = 0$ and $\alpha = 1$ hypotheses, in order to test the sensitivity to observe a long-term solar activity correlation. Each dataset was created by generating a random set of systematics ϵ according to a Gaussian distribution, scaling the expected number of events by the systematics, then generating events in each bin based on a Poisson distribution. An example of a toy dataset, showing the best-fit α solar correlation model, is shown in Figure 68. A large number of toys are generated and passed through the fitting procedure, and several tests are performed on the results to ensure that the fitting procedure functions as expected.

For example, for the $\alpha = 1$ sample, first the distribution of obtained α parameters is fit with a Gaussian function. The resulting parameters are mean $\mu = 1.03 \pm 0.02$ and width $\sigma = 0.57$, which confirms that the fit is unbiased. Similarly, the systematic error pull parameters fit with a mean consistent to zero, and the overall sample normalizations fit with a mean consistent to the nominal data rate used in the toy MC, confirming that the fitting of the systematic errors is also not biased. Next, the observed distribution of the likelihood ratio is fit with a χ^2 function, giving degrees-of-freedom $k = 96.0 \pm 0.4$, consistent with

---

4 In principle, the minimization over systematics can mean that Wilkes' theorem does not hold, and our test will not return a result distributed exactly as $\chi^2(1)$. However, tests using our toy MC (Sec. 7.2.6) showed that the impact of including the systematics seems to be negligible in the region of interest.

5 Which is somewhat analogous to the common case of $\sqrt{\chi^2}$.
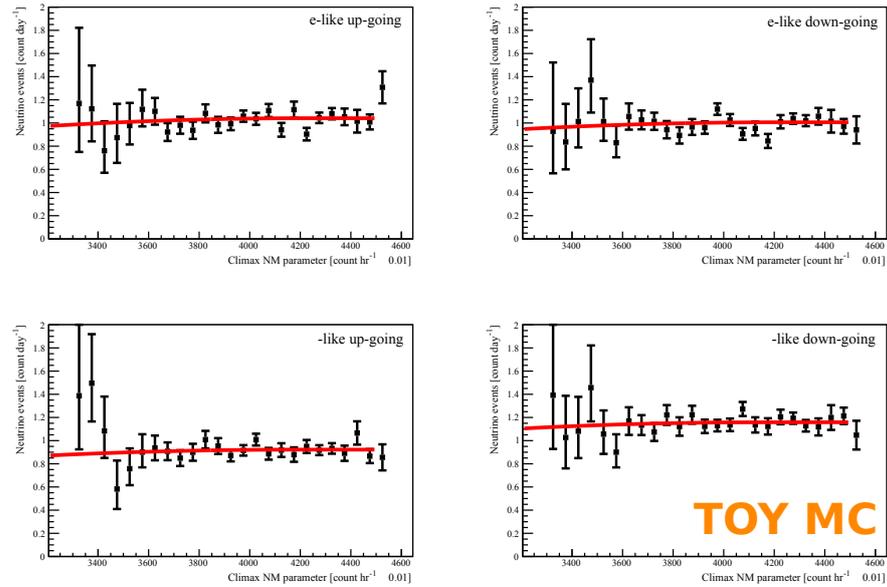
Figure 68: Example of a toy MC dataset, showing the four event type samples integrated over SK I-IV. To aid visualization (though contrary to the fitting procedure), here the best-fit solar correlation model is shown without being systematically shifted, and the data are systematically shifted and shown with Poisson error bars (using "Pearson's-χ^2 intervals" as defined in [143]).

96 free parameters (100 bins minus 4 sample normalizations). Finally, the significance $\sqrt{\Lambda}$ is plotted, and is distributed consistent with a $\chi^2(k = 1)$ as expected; by taking the mean of this distribution, we find that we expect on average a 1.75 σ sensitivity to observe a non-zero solar activity correlation.

7.2.7 *Search for effects during Forbush decreases*

The search for a long-term correlation between the neutrino flux and the solar activity, as described above, used events down to a Climax NM parameter of 3300 counts hr^{-1} ×0.01. In fact, the effective parameter sometimes drops significantly below this value, down to a minimum of ∼2800 counts hr^{-1} ×0.01, although the statistics are quite low and spread over a range of values. Such events are listed in Tab. 20; the cause of these events is believed to be large magnetically-charged coronal mass ejections, causing relatively short and high-intensity increases in the solar plasma flux [144]. On Earth these are associated with particularly large decreases in the cosmic ray flux and are often termed "Forbush decreases". While there is some ambiguity in the term, a common definition is any event that causes a > 10% decrease in the intensity of cosmic rays on the Earth's surface.

The reason for not using these time periods in the above analysis was the low statistics, and also that we suspect the correlation

START TIME (UTC)	END TIME (UTC)	OBSERVATION TIME [HR]
2000/07/15, 18:00	2000/07/17, 21:00	50
2001/04/11, 23:00	2001/04/13, 14:00	38
2003/10/29, 11:00	2003/11/01, 00:00	61
2003/11/01, 00:00	2003/11/04, 13:00	67
2005/01/19, 00:00	2005/01/19, 13:00	13
	TOTAL:	229

Table 20: The periods for which the Climax NM parameter drops below 3300, corresponding to strong Forbush decrease events.

between NMs at various rigidities and locations may not hold during Forbush decrease events. For these reasons, the HKKM simulation of the neutrino flux decrease is only modeled only down to $3500 \, \text{counts} \, \text{hr}^{-1} \times 0.01$, and although we extend the correlation down to $3300 \, \text{counts} \, \text{hr}^{-1} \times 0.01$ by a polynomial fitting, we believe extending it any further may be unrealistic.

Thus, while no theoretical prediction of the fractional neutrino flux decrease during these events is available, these may be the most sensitive times to measure a solar effect on the atmospheric neutrino flux. We thus make a simple test against the null hypothesis, by counting the number of events observed during all such periods over all four samples (in order to gain sufficient statistics), and comparing with the expected nominal event rate.

7.3 RESULTS AND DISCUSSIONS

7.3.1 *Continuous correlation search*

The data for the continuous correlation search, summing over SK I-IV and dividing into the four sample types (*e*-like or *μ*-like, and up-going or down-going) are shown in Fig. 69. The test statistic depending on α is shown in Fig. 70. The best-fit value of alpha is $\alpha = 0.62^{+(0.57, \, 1.12, \, 1.66)}_{-(0.58, \, 1.18, \, 1.80)}$, with errors given at (1 σ, 2 σ, 3 σ). The rejection power of the null hypothesis $\alpha = 0$ is 1.06 σ. The data in Fig. 69 seem to be in good agreement with the model, and the four plots together have a χ^2 statistic of 88.0 for 100 DOF, which is clearly reasonable.[6] The significance of rejecting the null hypothesis is lower than the mean value predicted by toy MC with $\alpha = 1$, but still a reasonably likely result according to that MC, at the p = 0.26 level. We show also in Fig. 71 the measured event rates binned per year, which is not used

---

6 The α-function based on the HKKM model however gives a similar goodness-of-fit to a simple linear fit, with $\chi^2 = 87.9$.
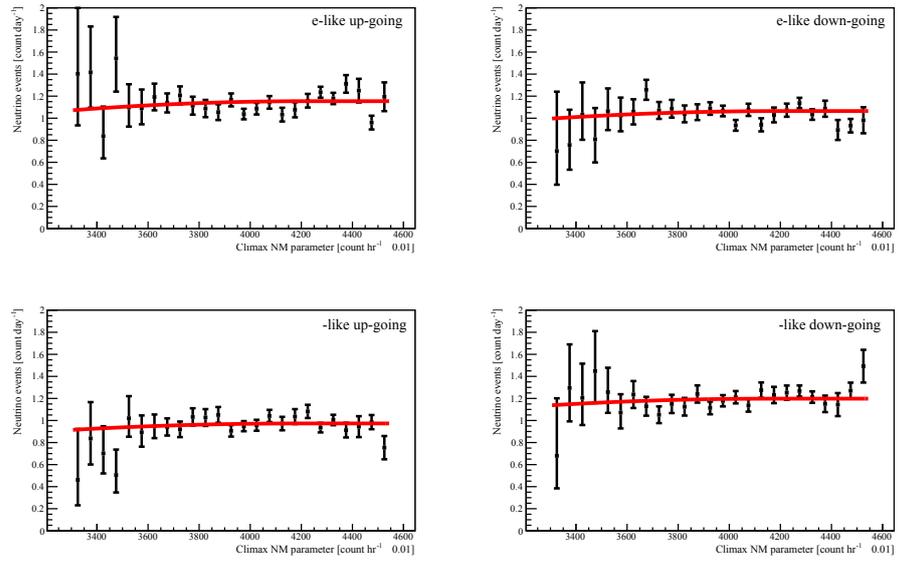
Figure 69: The test for a solar modulation correlation using SK-I to SK-IV data, showing the best-fit α-function (for $\alpha = 0.62$) across four data samples.
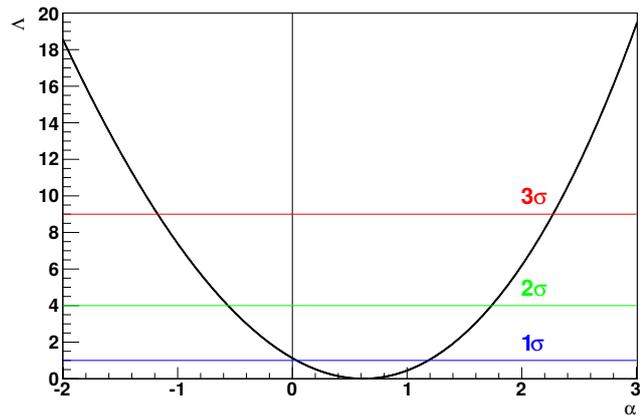
Figure 70: The test statistic, defined as $\Delta\chi^2$, as a function of α. The significance levels are drawn according to Wilks' theorem.
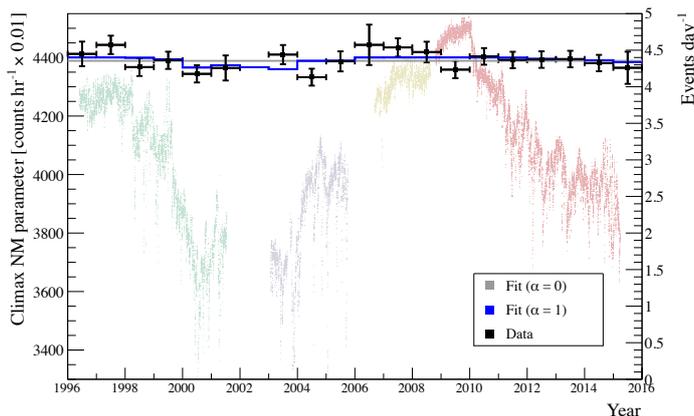
Figure 71: The measured daily event rate summing over the four selected Super-K data samples, compared against a linear function ($\alpha = 0$) and the predicted solar correlation function ($\alpha = 1$), where both functions are scaled to their best-fit normalizations. The NM count parameter, as in Fig. 63, is shown in the background for reference. Error bars represent statistical error. This plot is for illustrative purposes and is not used in the final analysis.

in the analysis, but gives an easily understandable demonstration of the sensitivity of our data.

We also perform the continuous correlation search on sub-samples of the data: first we re-run the analysis for each SK period I to IV separately, then we re-run the analysis for all SK periods but fitting using each sub-sample individually. The results are shown in Fig. 72. Data from the SK-III period alone cannot give any result, as it has no observation time above the minimum solar activity required to cause any effect according to the HKKM model. Although SK-II and SK-IV prefer a low value of α, the statistical power is lower and not inconsistent with the overall result.[7] Somewhat interestingly, the *e*-like samples prefer no correlation, while the μ-like samples prefer more the expected $\alpha = 1$ correlation; the significance is however not high enough to draw any strong conclusions.

Since the current HKKM model is used to apply solar activity weights in the SK MC, the strength of the correlation should be tested again by further measurements if possible, to ensure that the MC solar weighting does not bias the interpretation of the measurement of other neutrino parameters. For the Hyper-Kamiokande case, an estimation of the expected sensitivity to a long-term correlation depends strongly on the intensity and number of solar maxima observed in the future, which cannot be accurately predicted. However, if HK had observed the same time periods as SK, we calculate that a 5 σ (3 σ)

---

7 Indeed much of the fitting power comes from the *combination* of the data, and the overall result is thus not given by the average of these individual results.
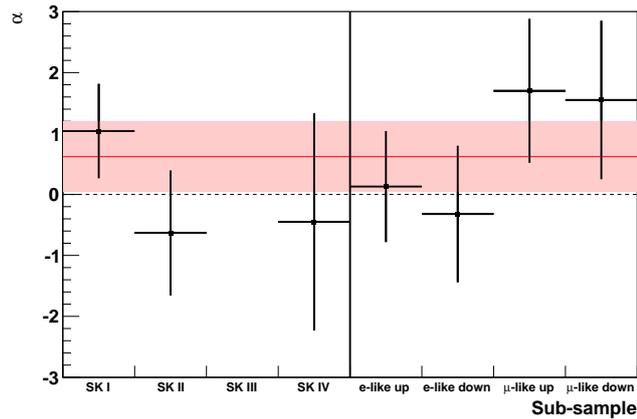
Figure 72: The best-fit α parameter obtained by performing the analysis on sub-samples of the data, showing SK periods on the left, and data sub-sample types on the right. The red band shows the combined result. The SK-III data point is absent, due to a lack of any data in the sensitive solar-active region, and thus a complete inability to perform the fit.

discovery sensitivity would have been expected for the true case of $\alpha = 1$ ($\alpha = 0.62$).

7.3.2 *Event rate during Forbush decreases*

During the coronal mass ejection periods described in Tab. 20, the SK detector was operational for a total of 7.21 days. Using the fitted normalization constants of the α-function in the previous analysis, which corresponds to the expected daily event rate with no solar modulation effect, we expect 31.80 ± 0.17 sub-GeV single-ring events in this time. The actual number of events recorded by the detector was $n_F = 20$, which by consulting the Poisson distribution gives $P(n_F \leq 20) = 0.017$, corresponding to rejection of the null hypothesis of no solar activity correlation at the 98.3% (2.38 σ) significance level.

This significance level is higher than the long-term correlation search, and while not high enough to claim evidence of a solar correlation, is nonetheless an indication that it may be possible to measure such an effect with high accuracy in the next generation neutrino detectors. In the Hyper-K case, based simply on the observed SK event counts, we would expect a 3 σ sensitivity from even a single Forbush decrease event.

CONCLUSIONS

Measurements of the atmospheric neutrino flux were performed using the water-Cherenkov detector Super-Kamiokande, which has the world's largest dataset of atmospheric neutrino interactions with an exposure of approximately $315\,\mathrm{kton\,yr}$. The atmospheric $\nu_e + \bar{\nu}_e$ and $\nu_\mu + \bar{\nu}_\mu$ fluxes were measured and analysed as a function of energy, azimuthal direction, and time, with a detailed analysis of systematic errors.

THE ENERGY SPECTRA were measured in the range $100\,\mathrm{MeV}$ up to $10\,\mathrm{TeV}$. Our measured data provided significantly improved precision up to $100\,\mathrm{GeV}$, and the first data below $320\,\mathrm{MeV}$. While none of the current generation flux models are strongly inconsistent with our data measurement, there was some preference for the HKKM11 model as the best-fit to our data. Energy spectra measurements were also obtained for the neutrino and antineutrino fluxes separately, using enriched neutrino and antineutrino samples; these were the first such direct measurements obtained for the atmospheric neutrino flux, and also showed agreement with the current flux models.

THE AZIMUTHAL SPECTRA showed an east-to-west dipole asymmetry in the neutrino flux, caused by the geomagnetic field, for both neutrino flavours at $> 5\,\sigma$. This represents the discovery of the effect for the ν_μ flux. The strength of this east-west effect was shown to depend on neutrino energy and zenith as predicted by the HKKM11 model. Furthermore, an indication of a predicted shift of the dipole asymmetry angle depending on the zenith angle was seen at the $2.2\,\sigma$ level, which is the first measurement that explores the geomagnetic field effects beyond a simple east-west asymmetry. These measurements give confidence that the neutrino flux simulations correctly model the complicated effects of the geomagnetic field.

A STUDY OF THE TIME CORRELATION between the atmospheric neutrino flux and the solar magnetic activity cycle was performed. The type of correlation predicted by the HKKM group was calculated to have a relatively minor effect on most of the Super-Kamiokande data, but by searching for a long-term correlation using approximately 14 years of data, a slight preference for such a correlation was seen at the $1.1\,\sigma$ level. By examining several much shorter periods (not included in the long-term analysis) corresponding to especially strong solar activity, from across the SK operational period for a total expo-

sure of 7.1 days, an indication for some decrease in the atmospheric neutrino flux was seen at the 2.4 σ level.

Summary and Future

These direct measurements of the atmospheric neutrino flux tested the theoretical models with improved precision, and searched for several new physical effects not previously measured. While these effects were seen at a range of significance levels, even those with relatively low significance are interesting indications, that may be further studied by the next generation of neutrino detectors. Our measurements are in general consistent with the current generation flux models, which gives confidence in our understanding and modeling of the atmospheric neutrino flux.

In future, combining our measurements with those from other current generation neutrino detectors, which are sensitive at distinct but overlapping energy regions, even more accurate data may be obtained. For example, constraints given by the improved precision of our energy spectra measurement combined with other measurements at higher energies can help to accurately determine the astrophysical neutrino spectra. In general, further improvements to the atmospheric flux measurements could also provide feedback to the flux simulations, and further a better systematic understanding of the atmospheric neutrino flux as both a background and a signal source.

Part IV

APPENDIX

A

SK SYSTEMATIC ERROR DATABASE

Here we list all systematic errors in the SK systematic error database, excluding some non-standard systematics described in the text (such as the azimuthal energy scale systematic, or the $\nu/\bar{\nu}$ separation uncertainties).

A.1 FLUX

The flux-related errors are based in the estimated uncertainties on the HKKM11 [19] atmospheric neutrino flux simulation. If numbers in brackets are given, they show the estimated errors at $1\,\sigma$.

- Absolute normalization at $E_\nu < 1$ GeV

- Absolute normalization at $E_\nu > 1$ GeV

- $\nu/\bar{\nu}$ ratio at $E_\nu < 1$ GeV (1%)

- $\nu/\bar{\nu}$ ratio at $1 < E_\nu < 10$ GeV (1.5%)

- $\nu/\bar{\nu}$ ratio at $E_\nu > 10$ GeV (2.5% until 30 GeV, then increases as $0.081 \times \log_{10}(E_\nu\,[\text{GeV}]) - 0.093$)

- $\nu_e/\bar{\nu}_e$ ratio at $E_\nu < 1$ GeV (2.5%)

- $\nu_e/\bar{\nu}_e$ ratio at $1 < E_\nu < 10$ GeV (2.5%)

- $\nu_e/\bar{\nu}_e$ ratio at $E_\nu > 10$ GeV (4% until 100 GeV, then increases as $0.067 \times \log_{10}(E_\nu\,[\text{GeV}]) - 0.103$)

- $\nu_\mu/\bar{\nu}_\mu$ ratio at $E_\nu < 1$ GeV (1%)

- $\nu_\mu/\bar{\nu}_\mu$ ratio at $1 < E_\nu < 10$ GeV (3%)

- $\nu_\mu/\bar{\nu}_\mu$ ratio at $E_\nu > 10$ GeV (3% until 50 GeV, then increases as $0.105 \times \log_{10}(E_\nu\,[\text{GeV}]) - 0.123$)

- Relative normalization of the FC multi-GeV sample (as depending on the flux uncertainties, not the detector systematics)

- Relative normalization of the PC-stopping and UPMU-stopping samples (as above)

- Up / down ratio

- Horizontal / vertical ratio

- K/π ratio (where K and π are the parent particles of the neutrinos)

- ν path length (zenith-angle dependant)

- Solar activity (compared to a fitted constant in each SK-I to SK-IV period)

A.2   CROSS SECTION

The cross-section related errors come from uncertainties in the NEUT predictions, and are explained in Sec. 4.5.1.

- Axial mass

- Axial mass for QE and single-pion interactions

- CCQE cross-section (energy dependant, by comparison to other models)

- CCQE $\nu/\bar{\nu}$ cross-section ratio (energy dependant)

- CCQE ν_e/ν_μ cross-section ratio (energy dependant)

- Single meson cross-section

- DIS cross-section

- DIS model differences

- Coherent π cross-section

- NC / CC ratio

- CCQE cross-section for ν_τ and $\bar{\nu}_\tau$

- $\pi^0/\pi^\pm$ ratio

- Single pion cross-section ratio for $\nu/\bar{\nu}$

- DIS q^2 at high W

- DIS q^2 at low W

A.3   DETECTOR

The detector errors come from uncertainties in the reconstruction abilities of the detector. The reduction and reconstruction errors are explained in Sec. 4. Generally, each of these errors is applied seperately for SK-I to SK-IV.

- Ring separation

- Single-ring PID

- Multi-ring PID

- Energy calibration (the overall energy scale of the entire detector)

- Up / down relative energy calibration

- FC reduction efficiency

- PC reduction efficiency

- UPMU reduction efficiency

- Fiducial volume cut efficiency

- FC / PC separation

- Non-ν_e background in e-like samples

- Non-ν_e background in multi-GeV single-ring e-like samples

- Non-ν_e background in multi-GeV multi-ring e-like samples

- Non-ν_μ background in μ-like samples

- Hadron simulation (in SKDetsim; responsible for NC contamination into single-ring events)

- UPMU reduction energy cut

- UPMU reduction path length cut

- UPMU background reduction of non-showering events

- UPMU background reduction of showering events

- UPMU background reduction of stopping events

- UPMU stopping / through-going separation

- UPMU showering / non-showering separation

- PC stopping / through-going separation (top cap)

- PC stopping / through-going separation (barrel)

- PC stopping / through-going separation (bottom cap)

- Two-ring π^0 sample normalization

- Polfit single-ring π^0 (moves samples between single-ring π^0 and single-ring e-like)

- Decay electron tagging efficiency

- Decay electrons from pions (the number of decay electrons able to be tagged in various hadronic events)

- Multi-GeV multi-ring other separation (where the "other" sample contains events not considered well-reconstructed enough to be placed into the other multi-ring samples)

- Multi-GeV multi-ring e-like $\nu_e / \bar{\nu}_e$ separation (from an older likelihood; not used in this thesis)

A.4 OSCILLATION

Systematic errors related to uncertainties on the neutrino oscillation parameters (not used in the oscillation analysis, or in the atmospheric neutrino flux energy spectra measurements in this thesis).

- θ_{12}

- θ_{13}

- θ_{23} (error size taken from the T2K experiment)

- Δm_{12}^2

- Δm_{23}^2 (error size taken from the T2K experiment)

- Matter effects

POISSON LIKELIHOOD METHOD

B.1 POISSON LIKELIHOOD RATIO

The Poisson distribution is given by

$$L(n|h) = \frac{h^n e^{-h}}{n!} \tag{70}$$

where h is the number of events predicted by a hypothesis H, and n is the number of observed events. Summing over all data bins i, the log-likelihood of one particular given set of data N is

$$\ln L(N|H) = \sum_i \ln \left(\frac{h_i^{n_i} e^{-h_i}}{n_i!} \right) \tag{71}$$

and using this likelihood, the ratio

$$\lambda_H = -2 \ln \left(\frac{L(N|H)}{L(N|N)} \right) \tag{72}$$

can be shown [145] to be approximately distributed as $\chi^2(NDF)$ in the limit $n \to \infty$. If the hypothesis is a function of j free parameters, then $NDF \approx i - j - 1$.

B.2 WILKS' THEOREM

The maximum likelihood estimator $L_M(N|H)$, for the hypothesis H, is defined as the case where $L(N|H)$ is maximised by modifying the parameters of H. By Wilks' theorem [146], if we consider an alternate hypothesis F that is a subset of H with a reduction in parameter space of k parameters, then the ratio

$$\Lambda = 2 \ln \frac{L_M(N|H)}{L_M(N|F)} \tag{73}$$

is asymptotically distributed as $\chi^2(k)$ in the case that F is true. This is an important result that allows us to perform a significance test between nested hypotheses, without the need for a toy MC study to understand the result.

Using the notation from Eq. (159), we can write

$$\Lambda = \lambda_F - \lambda_H \tag{74}$$

and for the Poisson case, since $\lambda \sim \chi^2(NDF)$, we can make an analogy of Λ to the commonly-used gaussian $\Delta\chi^2$ statistic. In particle physics, the statistic Λ is often written as $\Delta\chi^2$, even when a poisson likelihood method is being used.

[1] C. V. Achar et al. Detection of muons produced by cosmic ray neutrinos deep underground. *Phys. Lett.*, 18:196–199, 1965. (Cited on pages 3 and 7.)

[2] F. Reines, M. F. Crouch, T. L. Jenkins, W. R. Kropp, H. S. Gurr, G. R. Smith, J. P. F. Sellschop, and B. Meyer. Evidence for high-energy cosmic-ray neutrino interactions. *Phys. Rev. Lett.*, 15:429–433, Aug 1965. (Cited on pages 3 and 7.)

[3] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998. (Cited on pages 3 and 7.)

[4] Y. Ashie et al. A measurement of atmospheric neutrino oscillation parameters by super-kamiokande i. *Phys. Rev.*, D71:112005, 2005. (Cited on pages 3, 30, and 31.)

[5] R. Wendell et al. Atmospheric neutrino oscillation analysis with sub-leading effects in Super-Kamiokande I, II, and III. *Phys.Rev.*, D81:092004, 2010. (Cited on pages 3, 47, 55, and 63.)

[6] J. Hosaka et al. Solar neutrino measurements in super-Kamiokande-I. *Phys.Rev.*, D73:112001, 2006. (Cited on page 3.)

[7] B. Aharmim et al. Electron energy spectra, fluxes, and day-night asymmetries of B-8 solar neutrinos from measurements with NaCl dissolved in the heavy-water detector at the Sudbury Neutrino Observatory. *Phys.Rev.*, C72:055502, 2005. (Cited on page 3.)

[8] S. Abe et al. Precision Measurement of Neutrino Oscillation Parameters with KamLAND. *Phys.Rev.Lett.*, 100:221803, 2008. (Cited on pages 3 and 8.)

[9] F.P. An et al. Observation of electron-antineutrino disappearance at Daya Bay. *Phys.Rev.Lett.*, 108:171803, 2012. (Cited on pages 3 and 8.)

[10] J.K. Ahn et al. Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment. *Phys.Rev.Lett.*, 108:191802, 2012. (Cited on pages 3 and 8.)

[11] Y. Abe et al. Indication for the disappearance of reactor electron antineutrinos in the Double Chooz experiment. *Phys.Rev.Lett.*, 108:131801, 2012. (Cited on pages 3 and 8.)

[12] M.H. Ahn et al. Measurement of Neutrino Oscillation by the K2K Experiment. *Phys.Rev.*, D74:072003, 2006. (Cited on pages 3 and 8.)

[13] P. Adamson et al. Measurement of the neutrino mass splitting and flavor mixing by MINOS. *Phys.Rev.Lett.*, 106:181801, 2011. (Cited on pages 3, 8, and 83.)

[14] K. Abe et al. Precise Measurement of the Neutrino Mixing Parameter θ_{23} from Muon Neutrino Disappearance in an Off-Axis Beam. *Phys.Rev.Lett.*, 112(18):181801, 2014. (Cited on pages 3 and 8.)

[15] K. Abe et al. Observation of Electron Neutrino Appearance in a Muon Neutrino Beam. *Phys.Rev.Lett.*, 112:061802, 2014. (Cited on pages 3 and 8.)

[16] Paolo Lipari and Maurizio Lusignoli. Comparison of neutrino(mu) <—> neutrino(tau) and neutrino(mu) <—> neutrino(s) oscillations as solutions of the atmospheric neutrino problem. *Phys.Rev.*, D58:073005, 1998. (Cited on page 3.)

[17] K. Abe et al. Physics Potential of a Long Baseline Neutrino Oscillation Experiment Using J-PARC Neutrino Beam and Hyper-Kamiokande. *PTEP*, 2015. (Cited on pages 3 and 15.)

[18] Morihiro Honda, T. Kajita, K. Kasahara, S. Midorikawa, and T. Sanuki. Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data. *Phys.Rev.*, D75:043006, 2007. (Cited on pages 3, 23, 80, and 104.)

[19] M. Honda, T. Kajita, K. Kasahara, and S. Midorikawa. Improvement of low energy atmospheric neutrino flux calculation using the jam nuclear interaction model. *Phys. Rev. D*, 83:123001, Jun 2011. (Cited on pages 3, 23, 84, 87, 100, 103, 104, 106, 119, and 155.)

[20] G. Battistoni, A. Ferrari, T. Montaruli, and P.R. Sala. The fluka atmospheric neutrino flux calculation. *Astroparticle Physics*, 19(2):269 – 290, 2003. (Cited on pages 3, 85, and 104.)

[21] G. Barr, T. Gaisser, P. Lipari, S. Robbins, and T. Stanev. Three-dimensional calculation of atmospheric neutrinos. *Phys. Rev. D*, 70:023006, Jul 2004. (Cited on pages 3, 85, 104, and 133.)

[22] T. Sanuki, M. Motoki, H. Matsumoto, E.S. Seo, J.Z. Wang, et al. Precise measurement of cosmic ray proton and helium spectra with the BESS spectrometer. *Astrophys.J.*, 545:1135, 2000. (Cited on pages 3 and 23.)

[23] Sadakazu Haino, T. Sanuki, K. Abe, K. Anraku, Y. Asaoka, et al. Measurements of primary and atmospheric cosmic - ray spectra with the BESS-TeV spectrometer. *Phys.Lett.*, B594:35–46, 2004. (Cited on pages 3 and 23.)

[24] J. Alcaraz et al. Cosmic protons. *Phys.Lett.*, B490:27–35, 2000. (Cited on pages 3 and 23.)

[25] J. Alcaraz et al. Helium in near Earth orbit. *Phys.Lett.*, B494:193–202, 2000. (Cited on pages 3 and 23.)

[26] M.G. Catanesi et al. Measurement of the production cross-sections of $\pi^\pm$ in p-C and $\pi^\pm$-C interactions at 12 GeV/c. *Astropart.Phys.*, 29:257–281, 2008. (Cited on page 3.)

[27] M.G. Catanesi et al. Forward pi+/- production in p-O(2) and p-N(2) interactions at 12-GeV/c. *Astropart.Phys.*, 30:124–132, 2008. (Cited on page 3.)

[28] K. Daum. Determination of the atmospheric neutrino spectra with the frejus detector. *Z. Phys.*, C66:417–428, 1995. (Cited on pages 3 and 108.)

[29] R. Abbasi et al. Determination of the Atmospheric Neutrino Flux and Searches for New Physics with AMANDA-II. *Phys.Rev.*, D79:102005, 2009. (Cited on pages 4 and 108.)

[30] R. Abbasi et al. The Energy Spectrum of Atmospheric Neutrinos between 2 and 200 TeV with the AMANDA-II Detector. *Astropart.Phys.*, 34:48–58, 2010. (Cited on pages 4 and 108.)

[31] R. Abbasi et al. Measurement of the atmospheric neutrino energy spectrum from 100 gev to 400 tev with icecube. *Phys. Rev.*, D83:012001, 2011. (Cited on pages 4 and 108.)

[32] M. Aartsen et al. Measurement of the atmospheric ν_e flux in icecube. *Phys. Rev. Lett.*, 110:151105, Apr 2013. (Cited on pages 4 and 108.)

[33] R. Abbasi et al. A Search for a Diffuse Flux of Astrophysical Muon Neutrinos with the IceCube 40-String Detector. *Phys.Rev.*, D84:082001, 2011. (Cited on pages 4 and 108.)

[34] Chris Quigg. Cosmic Neutrinos. 2008. (Cited on page 5.)

[35] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957. (Cited on page 6.)

[36] Wolfgang Pauli. Open letter to the gauverein meeting in tubinge (available online). 1930. (Cited on page 7.)

[37] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. Detection of the free neutrino: a confirmation. *Science*, 124(3212):103–104, 1956. (Cited on page 7.)

[38] Jr. Davis, Raymond, Don S. Harmer, and Kenneth C. Hoffman. Search for neutrinos from the sun. *Phys.Rev.Lett.*, 20:1205–1209, 1968. (Cited on page 7.)

[39] Q.R. Ahmad et al. Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory. *Phys.Rev.Lett.*, 87:071301, 2001. (Cited on page 7.)

[40] K. S. Hirata et al. Observation in the kamiokande-ii detector of the neutrino burst from supernova sn1987a. *Phys. Rev. D*, 38:448–458, Jul 1988. (Cited on pages 8 and 14.)

[41] M.G. Aartsen et al. Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data. *Phys.Rev.Lett.*, 113:101101, 2014. (Cited on pages 8 and 20.)

[42] Z. Maki, M. Nakagawa, and S. Sakata. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, 28:870, 1962. (Cited on page 8.)

[43] B. Pontecorvo. Neutrino experiments and the question of leptonic-charge conservation. *Sov. Phys. JETP*, 26:984–988, 1968. (Cited on page 8.)

[44] Evgeny Kh. Akhmedov. Do charged leptons oscillate? *JHEP*, 0709:116, 2007. (Cited on page 9.)

[45] R.N. Mohapatra and P.B. Pal. Massive neutrinos in physics and astrophysics. Second edition. *World Sci.Lect.Notes Phys.*, 60:1–397, 1998. (Cited on page 9.)

[46] Boris Kayser. On the quantum mechanics of neutrino oscillation. *Phys. Rev. D*, 24:110–116, Jul 1981. (Cited on page 11.)

[47] C. Giunti, C.W. Kim, and U.W. Lee. When do neutrinos really oscillate?: Quantum mechanics of neutrino oscillations. *Phys.Rev.*, D44:3635–3640, 1991. (Cited on page 11.)

[48] C. Giunti, C.W. Kim, and U.W. Lee. Comments on the weak states of neutrinos. *Phys.Rev.*, D45:2414–2420, 1992. (Cited on page 11.)

[49] Daniel Kruppke. On Theories of Neutrino Oscillations: A Summary and Characterisation of the Problematic Aspects. 2007. (Cited on page 11.)

[50] L. Wolfenstein. Neutrino oscillations in matter. *Phys. Rev. D*, 17:2369–2374, May 1978. (Cited on pages 12 and 14.)

[51] Carlo Giunti and Chung W. Kim. Fundamentals of Neutrino Physics and Astrophysics. 2007. (Cited on pages 12 and 18.)

[52] T. K. Kuo and James Pantaleone. Neutrino oscillations in matter. *Rev. Mod. Phys.*, 61:937–979, Oct 1989. (Cited on page 12.)

[53] S.P. Mikheyev and A.Yu. Smirnov. Resonant amplification of Î‌œ oscillations in matter and solar-neutrino spectroscopy. *Il Nuovo Cimento C*, 9(1):17–26, 1986. (Cited on page 14.)

[54] H. A. Bethe. Possible explanation of the solar-neutrino puzzle. *Phys. Rev. Lett.*, 56:1305–1308, Mar 1986. (Cited on page 14.)

[55] Adapted from GNU free license data from http://commons.wikimedia.org. (Cited on pages 14, 29, and 31.)

[56] A.S. Barabash. Review of double beta decay experiments. 2014. (Cited on page 14.)

[57] Fanny Dufour and Dave Wark. Experimental Status of Neutrino Physics. *J.Phys.Conf.Ser.*, 408:012001, 2013. (Cited on page 15.)

[58] A. Renshaw et al. First Indication of Terrestrial Matter Effects on Solar Neutrino Oscillation. *Phys.Rev.Lett.*, 112(9):091805, 2014. (Cited on page 15.)

[59] K.A. Olive et al. Review of Particle Physics. *Chin.Phys.*, C38:090001, 2014. (Cited on pages 15, 84, 86, and 97.)

[60] Tsutomu Yanagida. Horizontal symmetry and masses of neutrinos. *Progress of Theoretical Physics*, 64(3):1103–1105, 1980. (Cited on page 17.)

[61] C. Giunti. Sterile Neutrino Status. 2013. (Cited on page 18.)

[62] Pasquale Blasi. The Origin of Galactic Cosmic Rays. 2013. (Cited on page 19.)

[63] M. Aguilar et al. First result from the alpha magnetic spectrometer on the international space station: Precision measurement of the positron fraction in primary cosmic rays of 0.5–350 gev. *Phys. Rev. Lett.*, 110:141102, Apr 2013. (Cited on page 19.)

[64] Rikard Enberg, Mary Hall Reno, and Ina Sarcevic. Prompt neutrino fluxes from atmospheric charm. *Phys.Rev.*, D78:043005, 2008. (Cited on page 20.)

[65] Image credit to the University of Delaware. (Cited on page 20.)

[66] Image credit to the Institute for Cosmic Ray Research. (Cited on pages 22, 32, 33, 57, 60, 61, 62, and 64.)

[67] T.K. Gaisser and M. Honda. Flux of atmospheric neutrinos. *Ann.Rev.Nucl.Part.Sci.*, 52:153–199, 2002. (Cited on page 22.)

[68] M. Honda, M. Sajjad Athar, T. Kajita, K. Kasahara, and S. Midorikawa. Atmospheric neutrino flux calculation using the NRLMSISE00 atmospheric model. 2015. (Cited on pages 23, 68, 124, and 135.)

[69] http://www.ngdc.noaa.gov/iaga/vmod/igrf.html. (Cited on page 24.)

[70] Tatsuhiko Sato et al. Particle and heavy ion transport code system PHITS version 2.52. *Journal of Nuclear Science and Technology*, 50(9):913–923, 2013. (Cited on page 24.)

[71] Stefan Roesler, Ralph Engel, and Johannes Ranft. The Monte Carlo event generator DPMJET-III. pages 1033–1038, 2000. (Cited on page 24.)

[72] J. M. Picone, A. E. Hedin, D. P. Drob, and A. C. Aikin. Nrlmsise-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *Journal of Geophysical Research: Space Physics*, 107(A12):SIA 15–1–SIA 15–16, 2002. (Cited on page 24.)

[73] U.S. Standard Atmosphere 1976. *U.S. Government Printing Office, Washington, D.C.*, 1976. (Cited on page 24.)

[74] Morihiro Honda. Private communication. (Cited on pages 25 and 136.)

[75] Morihiro Honda. Study of the Interaction Model Using Atmospheric Muons for the Calculation of Atmospheric Neutrino Flux. In *14th Workshop on Elastic and Diffractive Scattering (EDS Blois Workshop) on Frontiers of QCD (shown but not printed in the proceedings)*, volume C111215, 2011. (Cited on page 25.)

[76] Internal document, Institute for Cosmic Ray Research. (Cited on pages 31, 37, 52, 75, and 79.)

[77] Y. Fukuda et al. The Super-Kamiokande detector. *Nucl.Instrum.Meth.*, A501:418–462, 2003. (Cited on pages 32 and 42.)

[78] A. Suzuki, M. Mori, K. Kaneyuki, T. Tanimori, J. Takeuchi, et al. Improvement of 20-inch diameter photomultiplier tubes. *Nucl.Instrum.Meth.*, A329:299–313, 1993. (Cited on page 32.)

[79] K. Abe, Y. Hayato, T. Iida, K. Iyogi, J. Kameda, et al. Calibration of the Super-Kamiokande Detector. *Nucl.Instrum.Meth.*, A737:253–272, 2014. (Cited on pages 33, 36, 39, 40, 42, and 122.)

[80] H. Nishino, K. Awai, Y. Hayato, S. Nakayama, K. Okumura, M. Shiozawa, A. Takeda, K. Ishikawa, A. Minegishi, and Y. Arai. High-speed charge-to-time converter asic for the super-kamiokande detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 610(3):710 – 717, 2009. (Cited on page 34.)

[81] Giada Carminati. The new wide-band solar neutrino trigger for super-kamiokande. *Physics Procedia*, 61(0):666 – 672, 2015. 13th International Conference on Topics in Astroparticle and Underground Physics, 2013. (Cited on page 35.)

[82] S. Yamada et al. Commissioning of the new electronics and online system for the super-kamiokande experiment. *Nuclear Science, IEEE Transactions on*, 57(2):428–432, April 2010. (Cited on page 35.)

[83] H. Watanabe et al. First Study of Neutron Tagging with a Water Cherenkov Detector. *Astropart.Phys.*, 31:320–328, 2009. (Cited on pages 36 and 68.)

[84] Tristan James Irvine. PhD thesis, University of Tokyo. 2015. (Cited on pages 36, 68, 70, 71, 110, and 112.)

[85] Pietro Antonioli et al. Snews: the supernova early warning system. *New Journal of Physics*, 6(1):114, 2004. (Cited on page 37.)

[86] C. H. Townes, E. H. Wishnow, D. D. S. Hale, and B. Walp. A systematic change with time in the size of betelgeuse. *The Astrophysical Journal Letters*, 697(2):L127, 2009. (Cited on page 37.)

[87] http://www.caen.it/. (Cited on page 37.)

[88] http://iseg-hv.com/. (Cited on page 37.)

[89] Lee Ka Pik. PhD thesis, University of Tokyo. 2013. (Cited on pages 49, 55, 71, 73, and 79.)

[90] Chizue Ishihara. PhD thesis, University of Tokyo. 2010. (Cited on page 50.)

[91] Masato Shiozawa. PhD thesis, University of Tokyo. 1999. (Cited on pages 56 and 60.)

[92] T. Barszczak. PhD thesis, University of California, Irvine, 2005. (Cited on page 66.)

[93] John F. Beacom and Mark R. Vagins. GADZOOKS! Antineutrino spectroscopy with large water Cherenkov detectors. *Phys.Rev.Lett.*, 93:171101, 2004. (Cited on page 69.)

[94] M.R. Vagins. GADZOOKS! The Future of Super-Kamiokande? *Nuclear Physics B - Proceedings Supplements*, 168(0):128 – 130, 2007. Proceedings of the Neutrino Oscillation Workshop. (Cited on page 69.)

[95] M.T. Hussein, N.M. Hassan, and W. Elharbi. Hadron production in neutrino nucleon interactions at high-energies. *Int.J.Mod.Phys.*, A18:673, 2003. (Cited on page 70.)

[96] Fanny Dufour. PhD thesis, Boston University. 2009. (Cited on page 73.)

[97] Y. Ashie. et. al. Evidence for an oscillatory signature in atmospheric neutrino oscillations. *Phys. Rev. Lett.*, 93:101801, 2004. (Cited on page 76.)

[98] Eric Thrane. PhD thesis, University of Washington. 2008. (Cited on pages 76 and 80.)

[99] Shantanu Desai. PhD thesis, Boston University. 2004. (Cited on page 76.)

[100] S. Desai et al. Study of TeV neutrinos with upward showering muons in Super-Kamiokande. *Astropart.Phys.*, 29:42–54, 2008. (Cited on page 76.)

[101] M. E. C. Swanson and The Super-Kamiokande Collaboration. Search for diffuse astrophysical neutrino flux using ultra-high-energy upward-going muons in super-kamiokande i. *The Astrophysical Journal*, 652(1):206, 2006. (Cited on page 76.)

[102] The Super-Kamiokande Collaboration. High-energy neutrino astronomy using upward-going muons in super-kamiokande i. *The Astrophysical Journal*, 652(1):198, 2006. (Cited on page 76.)

[103] E. Thrane and The Super-Kamiokande Collaboration. Search for neutrinos from grb 080319b at super-kamiokande. *The Astrophysical Journal*, 697(1):730, 2009. (Cited on page 76.)

[104] E. Thrane and The Super-Kamiokande Collaboration. Search for astrophysical neutrino point sources at super-kamiokande. *The Astrophysical Journal*, 704(1):503, 2009. (Cited on page 76.)

[105] Yoshinari Hayato. A neutrino interaction simulation program library NEUT. *Acta Phys.Polon.*, B40:2477–2489, 2009. (Cited on page 80.)

[106] J.L. Hewett, H. Weerts, R. Brock, J.N. Butler, B.C.K. Casey, et al. Fundamental Physics at the Intensity Frontier. 2012. (Cited on page 81.)

[107] J.A. Formaggio and G.P. Zeller. From eV to EeV: Neutrino Cross Sections Across Energy Scales. *Rev.Mod.Phys.*, 84:1307, 2012. (Cited on pages 81 and 83.)

[108] C.H. Llewellyn Smith. Neutrino Reactions at Accelerator Energies. *Phys.Rept.*, 3:261–379, 1972. (Cited on page 82.)

[109] R.A. Smith and E.J. Moniz. NEUTRINO REACTIONS ON NUCLEAR TARGETS. *Nucl.Phys.*, B43:605, 1972. (Cited on page 82.)

[110] C.F. Perdrisat, V. Punjabi, and M. Vanderhaeghen. Nucleon Electromagnetic Form Factors. *Prog.Part.Nucl.Phys.*, 59:694–764, 2007. (Cited on page 82.)

[111] L. Fields et al. Measurement of Muon Antineutrino Quasielastic Scattering on a Hydrocarbon Target at $E\nu \sim$ 3.5 GeV. *Phys.Rev.Lett.*, 111(2):022501, 2013. (Cited on page 82.)

[112] R. Gran et al. Measurement of the quasi-elastic axial vector mass in neutrino-oxygen interactions. *Phys.Rev.*, D74:052002, 2006. (Cited on page 82.)

[113] J. Nieves, I. Ruiz Simo, and M.J. Vicente Vacas. The nucleon axial mass and the MiniBooNE Quasielastic Neutrino-Nucleus Scattering problem. *Phys.Lett.*, B707:72–75, 2012. (Cited on page 82.)

[114] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas. Inclusive charged-current neutrino-nucleus reactions. *Phys. Rev. C*, 83:045501, Apr 2011. (Cited on page 82.)

[115] Dieter Rein and Lalit M. Sehgal. Neutrino Excitation of Baryon Resonances and Single Pion Production. *Annals Phys.*, 133:79–153, 1981. (Cited on page 83.)

[116] S.K. Singh, M.J. Vicente-Vacas, and E. Oset. Nuclear effects in neutrinoproduction of Delta at intermediate energies. *Phys.Lett.*, B416:23–28, 1998. (Cited on page 83.)

[117] Dieter Rein and Lalit M. Sehgal. Coherent π^0 production in neutrino reactions. *Nuclear Physics B*, 223(1):29 – 44, 1983. (Cited on page 83.)

[118] M. Hasegawa et al. Search for coherent charged pion production in neutrino-carbon interactions. *Phys. Rev. Lett.*, 95:252301, Dec 2005. (Cited on page 83.)

[119] K. Hiraide et al. Search for charged current coherent pion production on carbon in a few-gev neutrino beam. *Phys. Rev. D*, 78:112004, Dec 2008. (Cited on page 83.)

[120] Y. Kurimoto et al. Improved measurement of neutral current coherent π^0 production on carbon in a few-gev neutrino beam. *Phys. Rev. D*, 81:111102, Jun 2010. (Cited on page 83.)

[121] M. Gluck, E. Reya, and A. Vogt. Dynamical parton distributions revisited. *The European Physical Journal C - Particles and Fields*, 5(3):461–470, 1998. (Cited on page 83.)

[122] A. Bodek and U.K. Yang. Modeling neutrino and electron scattering inelastic cross- sections in the few GeV region with effective LO PDFs TV Leading Order. *Nucl.Phys.Proc.Suppl.*, 2003. (Cited on page 83.)

[123] M. Derrick et al. Properties of the hadronic system resulting from $\bar{\nu}_\mu p$ interactions. *Phys. Rev. D*, 17:1–15, Jan 1978. (Cited on page 83.)

[124] S. Barlag et al. Charged hadron multiplicities in high energy $\bar{\nu}_\mu n$ and $\bar{\nu}_\mu p$ interactions. *Zeitschrift fur Physik C Particles and Fields*, 11(4):283–292, 1982. (Cited on page 83.)

[125] Torbjorn Sjostrand. PYTHIA 5.7 and JETSET 7.4: Physics and manual. 1995. (Cited on page 83.)

[126] Geant, cern program library long writeup w5013, march 1994. (Cited on page 84.)

[127] C. Zeitnitz and T.A. Gabriel. The GEANT - CALOR interface and benchmark calculations of ZEUS test calorimeters. *Nucl.Instrum.Meth.*, A349:106–111, 1994. (Cited on page 84.)

[128] K. Abe et al. Limits on Sterile Neutrino Mixing using Atmospheric Neutrinos in Super-Kamiokande. 2014. (Cited on page 84.)

[129] http://neutronm.bartol.udel.edu. (Cited on pages 85 and 138.)

[130] R. Wendell et al. Evidence for the appearance of atmospheric tau neutrinos in super-kamiokande. *Phys. Rev. Lett.*, 110:181802, May 2013. (Cited on page 91.)

[131] G. Cowan. A survey of unfolding methods for particle physics. *Conf.Proc.*, C0203181:248–257, 2002. (Cited on page 94.)

[132] G. D'Agostini. A multidimensional unfolding method based on bayes' theorem. *Nucl. Instrum. Meth. A*, 362(2-3):487 – 498, 1995. (Cited on page 94.)

[133] Tim Adye. Unfolding algorithms and tests using RooUnfold. pages 313–318, 2011. (Cited on pages 94 and 99.)

[134] G.D. Lafferty and T.R. Wyatt. Where to stick your data points: The treatment of measurements within wide bins. *Nucl.Instrum.Meth.*, A355:541–547, 1995. (Cited on page 97.)

[135] T. Futagami et al. Observation of the east - west anisotropy of the atmospheric neutrino flux. *Phys.Rev.Lett.*, 82:5194–5197, 1999. (Cited on page 119.)

[136] M. Yamada, K. Miyano, H. Miyata, H. Takei, M. Mori, et al. Measurements of the charge ratio and polarization of 1.2-TeV/c cosmic ray muons with the KAMIOKANDE-II detector. *Phys.Rev.*, D44:617–621, 1991. (Cited on page 124.)

[137] N.D. Gagunashvili. Pearson's chi-square test modifications for comparison of unweighted and weighted histograms and two weighted histograms. *PoS*, ACAT:060, 2007. (Cited on page 129.)

[138] Marius S. Potgieter. Solar modulation of cosmic rays. *Living Reviews in Solar Physics*, 10(3), 2013. (Cited on page 135.)

[139] JohnA. Simpson. The cosmic ray nucleonic component: The invention and scientific uses of the neutron monitor - (keynote lecture). *Space Science Reviews*, 93(1-2):11–32, 2000. (Cited on page 135.)

[140] http://cro.izmiran.rssi.ru/clmx/main.htm. (Cited on page 136.)

[141] http://www.nmdb.eu/. (Cited on pages 137 and 173.)

[142] http://cro.izmiran.rssi.ru/kerg/baseinfo.htm. (Cited on page 138.)

[143] Joel Heinrich. Coverage of error bars for poisson data. *CDF/MEMO/STATISTICS/PUBLIC/6438*, 2003. (Cited on page 146.)

[144] Arun Babu et al. High rigidity Forbush decreases: due to CMEs or shocks? 2013. (Cited on page 146.)

[145] Mark Thomson. Fitting and Hypothesis Testing (course notes, available from http://www.hep.phy.cam.ac.uk/ thomson/lectures/lectures.html). 2015. (Cited on page 159.)

[146] S. S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9(1):pp. 60–62, 1938. (Cited on page 159.)

[147] http://www.icrr.u-tokyo.ac.jp/ mhonda/. (Cited on page 173.)