An Improved Measurement of the ^{8}B Solar Neutrino Energy Spectrum at the Sudbury Neutrino Observatory

Gabriel D. Orebi Gann
Merton College, Oxford

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Trinity Term, 2008

An Improved Measurement of the ^{8}B Solar Neutrino Energy Spectrum at the Sudbury Neutrino Observatory

Gabriel D. Orebi Gann
Merton College, Oxford

Thesis submitted in partial fulfilment of the requirements for the
degree of Doctor of Philosophy at the University of Oxford
Trinity Term 2008

Abstract

The Sudbury Neutrino Observatory (SNO) has successfully demonstrated the phenomenon of neutrino flavour transformation in the flux of neutrinos from the Sun. Its unique use of heavy water as a detection medium allowed a measurement of both the total active solar neutrino flux, via the Neutral Current (NC) interaction, and the pure electron neutrino component via the Charged Current (CC) interaction. This thesis presents the most detailed analysis of the CC spectrum using the combined data from the first two phases of the experiment: the pure D_2O and the salt phases. The neutrino fluxes and spectra have been extracted down to an effective electron kinetic energy threshold of 3.5 MeV, lower than that achieved by any previous water Čerenkov experiment. The significant improvements made in the evaluation of the systematic uncertainties result in the most precise measurement of the total active solar neutrino flux ever reported:

$$\Phi_{\mathrm{NC}}^{uncon} = 5.191 \, ^{+0.156}_{-0.155}(\text{stat}) \, ^{+0.132}_{-0.099}(\text{syst}) \times 10^6 \, \text{cm}^{-2} \, \text{s}^{-1}$$

which is in good agreement with the flux predicted by the Standard Solar Model. The ratio of the CC and NC fluxes is found to be:

$$\frac{\Phi_{\mathrm{CC}}^{con}}{\Phi_{\mathrm{NC}}^{con}} = 0.312 \, ^{+0.013}_{-0.013}(\text{stat}) \, ^{+0.011}_{-0.010}(\text{syst})$$

which is a measure of the fraction of electron neutrinos in the total active neutrino flux from the Sun. The strong suppression observed is evidence for matter enhanced oscillations occurring in the Sun, as predicted by the MSW effect.

In addition to the observed overall ν_e flux suppression, a two-dimensional test of the MSW prediction for the spectral shape has been performed in order to extract the most likely values for the neutrino oscillation parameters. The resulting values are:

$$\Delta m_{12}^2 = 6.03 \times 10^{-5} \, \text{eV}^2 \qquad \tan^2 \theta_{12} = 0.42$$

with reduced uncertainties in the value of $\tan^2 \theta_{12}$ in comparison to previous SNO results. The addition of data from other experiments, such as KamLAND, should enable the most accurate determination of both these mixing parameters.

"Though this be madness, yet there is method in't"
Hamlet 2.2.193-206 – W. Shakespeare

Acknowledgements

Contents

List of Figures

List of Tables

Glossary

Acceptance Fraction of signal events passing a cut.

ADC Analogue to Digital Converter.

AMB Analogue Measurement Board.

AV Acrylic Vessel.

CC Charged Current interaction.

DAQ Data Aquisition system.

ECA Electronics Calibration.

ES Elastic Scattering interaction.

Data cleaning cuts Background rejection criteria applied to the data set.

FEC Front End Card.

FTK an energy reconstruction algorithm used in SNOMAN.

FTP Path Fitter - a position reconstruction algorithm used in SNOMAN.

FTU A position reconstruction algorithm used in SNOMAN.

GT Global Trigger.

GTID Global Trigger Identification number.

ITR In Time Ratio.

Julian Date Time scale used by SNO, measured from 00:00:00 GMT on 01/01/75.

LETA Low Energy Threshold Analysis.

LMA The Large Mixing Angle oscillation scenario.

MC Monte Carlo, used to refer to the Monte Carlo simulation code.

MCE Modified Collection Efficiency - an empirical function in the PMT simulation.

MINUIT A standard minimisation package from the CERN software library.

MSW The Mikheyev-Smirnov-Wolfenstein matter enhancement of neutrino oscillation.

MTC/A Analogue Master Trigger Card.

MTC/D Digital Master Trigger Card.

MXF A maximum likelihood signal extraction software package.

NC Neutral Current interaction.

NCD Neutral Current Detector.

NHit Number of PMTs recording a successful hit within a fixed time window.

NLL Negative Log Likelihood.

OWL Outward Looking PMT

PCA PMT Calibration.

PDF Probability Density Function.

PMT Photo Multiplier Tube.

PMTIC PMT Interface Card.

PSUP PMT Support structure.

RSP an energy reconstruction algorithm used in SNOMAN.

Sacrifice Fraction of signal events failing a cut.

SNO The Sudbury Neutrino Observatory.

SNOMAN The SNO Monte Carlo and Analysis software package.

SSM Standard Solar Model.

TAC Time-to-Amplitude Converter.

Chapter 1

The Theory of the Neutrino

Ever since its conception by Pauli in 1930 [1] to explain the continuous spectrum of nuclear beta decay, the neutrino has provided a wealth of controversial predictions. Originally thought to be undetectable, the first experimental observation was made by Reines and Cowan in 1956, utilising the delayed coincidence signal from inverse beta-decay [2]. The neutrino was thought to be massless, and was incorporated into the Standard Model of particle physics as such, until observations of neutrino flavour oscillation in the 1990s necessitated the introduction of a non-zero neutrino mass. Discrepancies between experimental measurements and of observation from theoretical prediction have led to a continuous evolution in the understanding of this unique particle. The aim of this thesis is to take this understanding one step further.

This chapter introduces the neutrino and its properties, including the theory behind neutrino flavour oscillation. The major experimental observations of this phenomenon are described, with focus on the solar neutrino experiments that led to the well known solar neutrino problem. The resolution of this problem by the second generation of solar neutrino experiments is presented, along with the as yet unanswered questions. The importance of a low energy measurement of the solar neutrino spectrum in completing the story is discussed. Such a measurement is the focus of this thesis.

1.1 The Physics of Neutrinos

The Standard Model is the theoretical framework that describes the fundamental particles of nature. Spin-half fermions are separated into the 'coloured' quarks and the colour-neutral leptons, which are further organised into three increasingly massive generations. Integer spin bosons mediate the three fundamental forces by which these particles interact: the strong force is mediated by gluons, the electromagnetic force by photons and the weak force by the $W^{\pm}$ and Z^0 bosons.

Incorporated into the Standard Model as a massless particle, the neutrino is the fundamental, neutral partner of the charged lepton. The three generations of neutrino are associated respectively with the three charged leptons: the electron (e), the muon (μ) and the tau (τ). All three of these *flavours* have been experimentally observed via their interactions with matter.

Experimental observations have been consistent with the neutrino having negative helicity (spin anti-parallel to its momentum). Therefore, it exists in the Standard Model in a left-handed doublet along with its charged lepton partner, with the right-handed field having a magnitude of zero. Similarly, the antimatter partner of the neutrino, the antineutrino, has positive helicity.

The Standard Model does not include a mechanism by which particles obtain mass, but the current theory involves coupling to the Higgs boson. This coupling requires non-zero terms for both the left and right-handed fields for each particle and therefore, with no right-handed field, the neutrino remains massless. Although the Standard Model does not allow for a non-zero neutrino mass, strong experimental evidence to the contrary necessitated its inclusion in the theory. Therefore, the Standard Model had to be extended to include the existence of a right-handed neutrino field.

The neutrino carries neither electrical nor colour charge and can therefore only interact via the aptly named *weak* interaction (massive neutrinos are also affected by the gravitational force, but this is not incorporated into the Standard Model). The weak interaction is mediated by the massive $W^{\pm}$ and Z^0 bosons. Measurements of the width of the Z^0 boson at CERN have determined the number of active massive neutrino states with masses less than half that of the Z^0 to be three, in agreement with the Standard Model prediction [3].

1.1.1   Neutrino Oscillation

Neutrinos are produced via the weak interaction in one of the three flavour eigenstates. Their speed of propagation is determined by their mass: they propagate as mass eigenstates. In an analogous manner to mixing in the quark sector, these mass eigenstates do not have to be identical to the flavour eigenstates. If the two sets of eigenstates are rotated with respect to one another then as the neutrino propagates, the relative phase of the flavour eigenstates increases monotonically, changing the composition of the particle state. This results in a finite probability that a neutrino created in one flavour state can be detected in a different flavour state after having propagated some finite distance. This phenomenon is known as neutrino oscillation [4].

The Maki-Nakagawa-Sakata-Pontecorvo (MNSP) mixing matrix describes the relation between the mass and flavour eigenstates:

$$|\nu_x\rangle = U_{xi} |\nu_i\rangle \tag{1.1}$$

where x represents the three flavour states (e, μ, τ) and i represents the three mass states (1, 2, 3). This matrix can be parameterised into three sectors as follows:

$$U_{xi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.2}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$, θ_{ij} is the mixing angle for each sector and δ is a CP violating phase. Since θ_{13} is small (see section 1.2.3) and $\Delta m_{12}^2 \ll \Delta m_{23}^2$, the central matrix reduces to the identity matrix and mixing in the remaining sectors can be approximated by two flavour oscillation. For the θ_{12} sector, the mixing is therefore parameterised by:

$$\begin{bmatrix} |\nu_e\rangle \\ |\nu_x\rangle \end{bmatrix} = \begin{bmatrix} \cos\theta_{ij} & \sin\theta_{ij} \\ -\sin\theta_{ij} & \cos\theta_{ij} \end{bmatrix} \begin{bmatrix} |\nu_1\rangle \\ |\nu_2\rangle \end{bmatrix} \tag{1.3}$$

where ν_x is an admixture of the ν_μ and ν_τ flavour states. This can be used to calculate the energy-dependent survival probability of an electron neutrino, which is the probability that it will be detected in the same flavour state after having travelled a finite distance:

$$P_{\nu_e \to \nu_e} = 1 - \sin^2 2\theta_{12} \sin^2\left(\frac{\Delta m_{12}^2 L}{4E}\right) \tag{1.4}$$

$$\Delta m_{12}^2 = m_2^2 - m_1^2 \tag{1.5}$$

using natural units ($\hbar = c = 1$) where E is the energy of the neutrino in GeV, L is the distance travelled in km and Δm_{12}^2 is the difference of the squared masses of the two mass eigenstates, ν_1 and ν_2, in eV^2, as given in equation 1.5. If the mass of the neutrino is zero, or if the mass difference between states is zero, then no flavour transformation will occur.

The oscillation length is defined as the distance over which a neutrino returns to its initial state. It follows from equation 1.4 that this is given by:

$$L_V = \frac{4\pi E}{\Delta m_{12}^2} \tag{1.6}$$

where L_V is the vacuum oscillation length. The survival probability can then be written as:

$$P_{\nu_e \to \nu_e} = 1 - \sin^2 2\theta_{12} \sin^2 \left(\frac{\pi L}{L_V} \right) \tag{1.7}$$

The objective of many experimental programmes is to measure the values of the two oscillation parameters, θ_{12} and Δm_{12}^2. The sensitivity of a particular experiment to these oscillations is governed by the neutrino propagation distance relative to the oscillation length. If $L \ll L_V$ then the oscillation will not have time to develop; the θ_{12} term in equation 1.7 can be neglected and the survival probability is essentially equal to 1.0. Alternatively, if the variation in the distance over which the neutrino travels between creation and detection is greater than the oscillation length ($\delta L \gg L_V$) then only the average of the oscillations is observed. This applies to neutrinos produced in an extended source, such as those generated in the core of the Sun. The oscillatory term in equation 1.7 takes its average value of 0.5, and the survival probability is then given by:

$$P_{\nu_e \to \nu_e} = 1 - \frac{1}{2} \sin^2 2\theta_{12} \tag{1.8}$$

1.1.2 Interactions with Matter

When neutrinos propagate through matter, additional interactions can complicate the simple vacuum oscillation scenario described above. All three flavours of neutrino experience coherent, elastic forward scattering interactions with the matter through which they pass. ν_μ and ν_τ can interact with the constituent electrons by Neutral Current (NC) processes only whereas ν_e can also interact via the Charged Current (CC) process. This results in an enhanced forward scattering amplitude for ν_e, changing the relative propagation of and thus the superposition of the three flavour states. Postulated by Wolfenstein [5] and applied to the Sun by Mikheyev and Smirnov [6], this "MSW" effect can result in large mixing even for small values of the vacuum oscillation mixing angle.

The inclusion of the CC scattering interactions experienced by ν_e results in an additional effective potential in the wave equation for ν_e, $V = \sqrt{2} G_F N_e$, where G_F is the Fermi constant and N_e is the number density of electrons in the medium. This adds to a diagonal term in the Hamiltonian governing the propagation of the vacuum mass eigenstates through matter, which results in an increased effective mass for ν_e.

The form for neutrino oscillations remains the same as that described previously, with modified forms for the mixing angle and the oscillation length. The mixing angle in matter is given by:

$$\sin^2 2\theta_m = \frac{\sin^2 2\theta_V}{(\cos 2\theta_V - \frac{L_V}{L_e})^2 + \sin^2 2\theta_V} \tag{1.9}$$

where θ_V is the vacuum mixing angle, L_V is the vacuum oscillation length as defined in equation 1.6 and L_e is the electron scattering length, defined by:

$$L_e = \frac{\sqrt{2}\pi}{G_F N_e} \tag{1.10}$$

From equation 1.9 it can be seen that θ_m takes a value of $\frac{\pi}{4}$ and, hence, maximal mixing will occur, even for small values of the vacuum mixing angle, when the electron density in the medium is such that $L_V = L_e \cos 2\theta_V$. This resonance condition is met when:

$$N_e = \frac{\Delta m_{12}^2 \cos 2\theta_V}{2\sqrt{2}EG_F} \tag{1.11}$$

This therefore imposes a condition on the neutrino energy, dependent on both the electron density in the medium and the values of the vacuum oscillation parameters, in order for resonance to occur.

Constant Electron Density

For a medium with a constant electron density, the mixing is constant in time with an oscillatory nature, as for the vacuum oscillation scenario. However, the oscillation parameters are modified and have an energy dependence, as described above. Therefore, matter oscillations would lead to a characteristic modification of the neutrino energy spectrum relative to the vacuum scenario. This situation can be applied to neutrinos crossing through the mantle of the Earth and the effect can be probed by comparing flux measurements during the day to those taken during the night.

Variable Electron Density

For a medium with a variable electron density, such as the Sun, the value of N_e is effectively changing with time as the neutrino propagates through the medium. The value of the matter mixing angle is therefore also time-dependent. As the neutrino travels through the varying density medium, there is a layer at which resonance occurs. It is within this layer that flavour transformation mainly occurs.

By substituting the matter mixing angle, θ_m, for the vacuum angle, θ_{12}, and considering mixing in terms of the matter eigenstates, ν_{1m} and ν_{2m}, equation 1.3 can be applied to the matter-mixing scenario. The composition of the neutrino mass eigenstates in matter is therefore given by:

$$|\nu_{1m}\rangle = \quad \cos\theta_m |\nu_e\rangle \quad + \sin\theta_m |\nu_x\rangle \tag{1.12}$$

$$|\nu_{2m}\rangle = \quad -\sin\theta_m |\nu_e\rangle \quad + \cos\theta_m |\nu_x\rangle \tag{1.13}$$

Neutrinos are created in the core of the sun, where the electron density is at its highest. For regions of very high N_e, $\sin^2 2\theta_m \sim 0$ (equation 1.9) and, hence, $\theta_m \sim \frac{\pi}{2}$. From equations 1.12 and 1.13, this corresponds to the electron neutrino consisting predominantly of the more massive state, ν_{2m}.

As the neutrino propagates outwards, the density of the Sun is falling, thus the mixing angle decreases and the amount of mixing increases. At a certain point, if the neutrino energy is sufficient, the resonance condition is met and maximal mixing occurs. As a result, the ν_{2m} state picks up a significant admixture of ν_x.

If the density of the Sun is changing slowly enough, then the adiabacity condition is met and transitions between the two matter eigenstates are suppressed, such that they can be considered as propagating independently. The neutrino therefore exits the Sun almost entirely in the ν_2 mass state, with an enhanced ν_x component. If this condition does not hold, then there is a small probability that transition could occur between the ν_{2m} and ν_{1m} states and the neutrino could exit in the ν_1 state, with a correspondingly lower proportion of ν_x. This admixture of ν_x in the final neutrino state results in a measured suppression of the ν_e survival probability, for either adiabatic or quasi-adiabatic conditions, relative to the vacuum oscillation scenario.

1.1.3 MSW Predictions

Spectral Distortion

For low energy neutrinos, the resonance condition is never met and the oscillation is governed solely by the vacuum oscillation scenario (equation 1.7). The survival probability of neutrinos with higher values of energy, such that resonance occurs within the Sun, is suppressed due to the MSW effect. This leads to an energy dependence of the ν_e component of the solar neutrino flux.

The MSW effect itself is also a strong function of energy, due to the energy dependence of the resonance condition. For certain values of the vacuum oscillation parameters, this can lead to distortions in the ^{8}B neutrino spectrum. In particular, a small distortion is predicted across the transition between vacuum and matter-dominated oscillation, which occurs in the energy range of $\sim 1 - 5\,\text{MeV}$. A solar neutrino experiment that was sensitive to these energies could therefore probe this effect.

Day-Night Effect

Matter oscillations can also occur within the Earth. Solar neutrinos detected during the day pass through very little of the Earth's matter in comparison to those detected at night. The MSW effect should, therefore, result in an asymmetry between the day and night fluxes. In the reverse of the process occurring in the Sun, ν_e would be regenerated in the Earth's matter, resulting in a larger solar ν_e flux at night than during the day. In the allowed region of parameter space only a small asymmetry is expected, of less than 5%. A high degree of experimental accuracy would be required to observe this effect.

1.2 Experimental Evidence for Neutrino Oscillation

The weakly interacting nature of the neutrino necessitates an intense source for any experimental measurement. The source employed affects the sensitivity of an experiment to the three sectors of neutrino oscillation, as defined in equation 1.2.

1.2.1 Cosmic Rays

Cosmic rays interact in the Earth's upper atmosphere, producing a flux of atmospheric neutrinos, which can be detected on the surface. Pion decay and the subsequent muon decay produce both ν_e and ν_μ, in a ratio of approximately 1:2:

$$\pi^- \rightarrow \mu^- \bar{\nu}_\mu$$
$$\mu^- \rightarrow e^- \nu_\mu \bar{\nu}_e$$

The neutrino energies are on the order of GeV and so atmospheric neutrino experiments are sensitive to the $[\theta_{23}, \Delta m_{23}^2]$ sector of neutrino oscillation. By observing the ratio of the two flavours of neutrino reaching the Earth's surface, a measurement can be made of the relevant oscillation parameters.

Super-Kamiokande (SuperK) is a water Čerenkov detector, which detects both ν_e and ν_μ by their charged current interactions in H_2O. Neutrino flavour identification is possible by identifying the resulting charged lepton. SuperK observed a strong zenith angle dependence to the ratio of detected ν_e and ν_μ, with a deficit of upward-going ν_μ but a roughly constant flux of ν_e, independent of direction [7]. This was interpreted as evidence for $\nu_\mu \rightarrow \nu_\tau$ oscillation, and thus proof that ν_μ had mass, and allowed SuperK to place tight limits on the oscillation parameters [8] of:

$$\sin^2 2\theta_{23} > 0.92 \qquad 1.5 < |\Delta m_{23}^2| < 3.4 \times 10^{-3} \, \text{eV}^2 \qquad (90\% \text{ C.L.}) \qquad (1.14)$$

1.2.2 Accelerators

Long baseline accelerator experiments are also sensitive to the atmospheric neutrino sector. High energy muon-neutrino beams, generated by colliding protons on a stationary target, are directed towards a far detector. The K2K [9] and MINOS [10] experiments have observed deficits of ν_μ in agreement with the oscillation parameters measured by SuperK. The current limits from MINOS are:

$$\sin^2 2\theta_{23} > 0.84 \qquad \Delta m_{23}^2 = 2.38^{+0.20}_{-0.16} \times 10^{-3} \, \text{eV}^2 \qquad (90\%, \ 68\% \ \text{C.L.}) \qquad (1.15)$$

The T2K experiment in Japan will have a significantly increased sensitivity to oscillation, allowing a more precise measurement of both the mixing angle and the mass difference in the atmospheric sector, as well as a measurement of θ_{13}.

1.2.3 Nuclear Reactors

Reactor experiments detect the flux of $\bar{\nu}_e$ generated from the beta decay of nuclear fission products. The CHOOZ experiment used a liquid scintillator target to detect neutrinos at a distance of $\sim 1\,\text{km}$ from the reactor. The neutrinos were detected by the inverse beta decay reaction with a target proton:

$$\bar{\nu}_e + p \rightarrow n + e^+$$
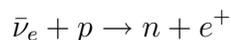
The delayed coincidence signal from the prompt positron and the delayed scintillation light from neutron capture results in low contamination from background 'singles' events. Reactor experiments are sensitive to the $[\theta_{13}, \ \Delta m_{13}^2]$ oscillation sector. CHOOZ observed no deficit in the expected $\bar{\nu}_e$ flux, implying that $\bar{\nu}_e$ is not strongly mixing with $\bar{\nu}_3$ [11]. Therefore, an upper bound was placed on the value of the relevant mixing angle of:

$$\sin^2 2\theta_{13} < 0.17 \qquad (90\% \ \text{C.L.}) \qquad (1.16)$$

1.2.4 The Sun

The Sun is an intense generator of electron neutrinos. Many experiments have utilised this naturally abundant source in order to observe the interactions of neutrinos. Solar neutrino experiments can be separated into two main categories, according to the method of particle detection: radiochemical and water Čerenkov. These are described in more detail in section 1.4. A deficit of ν_e in comparison to the predicted flux from theoretical

solar models was observed in both types of experiment, implying some level of oscillation of ν_e to ν_x. The resulting best-fit values for the relevant oscillation parameters [12] are:

$$\tan^2 \theta_{12} = 0.45\,^{+0.09}_{-0.04} \qquad \Delta m^2_{12} = 8.0\,^{+0.6}_{-0.4} \times 10^{-5}\,\mathrm{eV}^2 \qquad (1.17)$$

The solar neutrino experiments are the focus of the remainder of this chapter.

1.3 The Standard Solar Model

Thermonuclear reactions in the core of the Sun produce a large flux of electron neutrinos. The primary source of solar energy is the fusion process:

$$4p \rightarrow\,^4\mathrm{He} + 2\nu_e + 2e^+ \qquad (1.18)$$

which proceeds via a complex series of nuclear reactions known as the pp chain. Approximately 26.7 MeV is released in this reaction, with a small fraction being carried away as the kinetic energy imparted to the outgoing particles and the majority being released in the form of γ-rays. The full pp chain is shown in figure 1.1. Neutrinos are produced in five reactions in this chain, the energies of which are well known from nuclear physics experiments, independently of any theoretical solar models. These reactions result in the spectra of neutrino energies shown in figure 1.2. Neutrinos are also produced in the CNO chain, but these constitute a negligible proportion of the total solar neutrino flux ($< 1\%$ [13]) and so have been omitted from the discussion here.

A number of models have been developed to describe and predict the behaviour of the Sun. Referred to as Standard Solar Models (SSMs), these models are based on the premise that the Sun is a spherical body, held in equilibrium by the balance of radiative pressure and gravitational force, thus preventing its collapse. One such model is that devised by John Bahcall [13]: BS2005(OP). For the remainder of this thesis, the term 'SSM' is used to refer to this model. The SSM has made many successful predictions. In particular, the predictions for the speed of sound agreed with measurements made using helioseismology to a very high level of accuracy.

The SSM predicts the density profile of the Sun, which is critical in studying the impact of the MSW oscillations described in section 1.1.2. It also predicts the flux of neutrinos produced by each of the reactions in the pp chain. The size of the solar neutrino flux is well constrained by the precise knowledge of the Sun's luminosity, resulting in small theoretical uncertainties on the dominant neutrino fluxes. Knowledge of the nuclear cross-sections involved limits the accuracy of the predictions for the weaker branches. The flux predictions are summarised in table 1.1 along with the theoretical uncertainties from the

$(\boldsymbol{pp})$ $\boxed{p + p \to {}^2\mathrm{H} + e^+ + \boldsymbol{\nu}_e}$ $\qquad$ $\boxed{p + e^- + p \to {}^2\mathrm{H} + \boldsymbol{\nu}_e}$ $(\boldsymbol{pep})$

99.6% $\qquad\qquad$ 0.4%

$\boxed{{}^2\mathrm{H} + p \to {}^3\mathrm{He} + \gamma}$

85% $\qquad\qquad$ $2 \times 10^{-5}\,\%$

$\boxed{{}^3\mathrm{He} + {}^3\mathrm{He} \to {}^4\mathrm{He} + 2\,p}$ $\qquad\qquad$ $\boxed{{}^3\mathrm{He} + p \to {}^4\mathrm{He} + e^+ + \boldsymbol{\nu}_e}$

$(\boldsymbol{hep})$

15%

$\boxed{{}^3\mathrm{He} + {}^4\mathrm{He} \to {}^7\mathrm{Be} + \gamma}$

99.87% $\qquad\qquad$ 0.13%

$(\boldsymbol{{}^7\mathbf{Be}})$ $\boxed{{}^7\mathrm{Be} + e^- \to {}^7\mathrm{Li} + \boldsymbol{\nu}_e}$ $\qquad$ $\boxed{{}^7\mathrm{Be} + p \to {}^8\mathrm{B} + \gamma}$

$\boxed{{}^7\mathrm{Li} + p \to 2\,{}^4\mathrm{He}}$ $\qquad\qquad$ $\boxed{{}^8\mathrm{B} \to {}^8\mathrm{Be}^* + e^+ + \boldsymbol{\nu}_e}$ $(\boldsymbol{{}^8\mathbf{B}})$

$\boxed{{}^8\mathrm{Be}^* \to 2\,{}^4\mathrm{He}}$
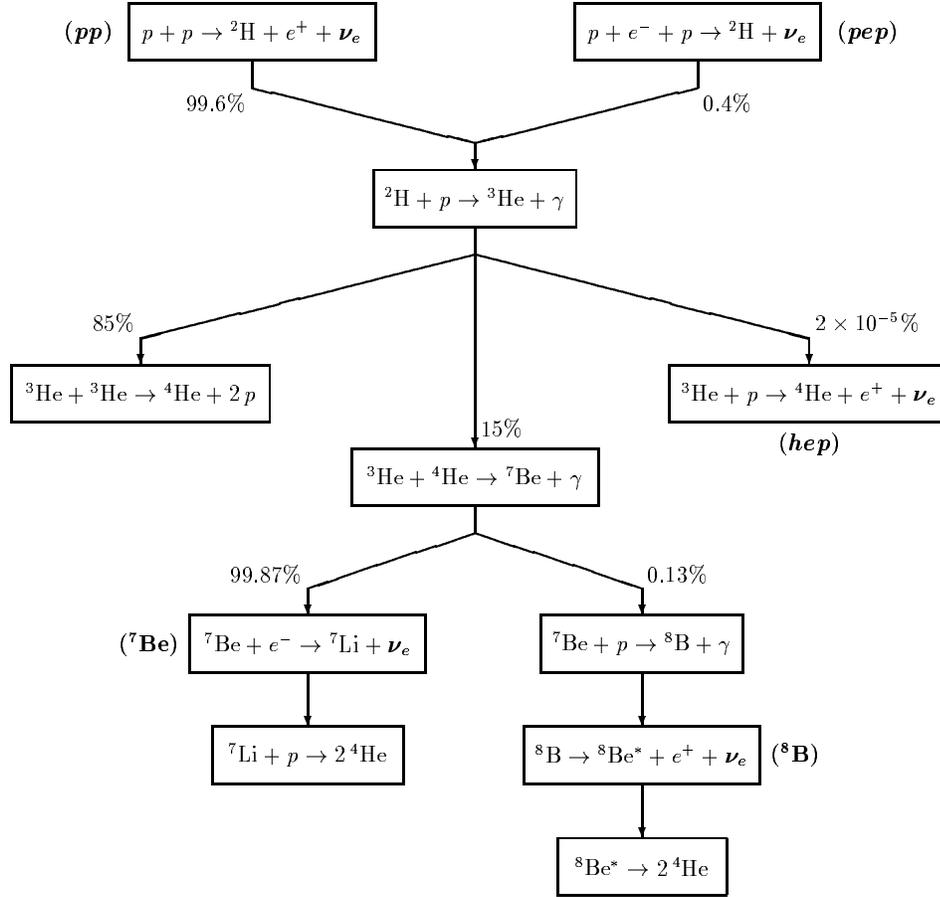
Figure 1.1: The nuclear reactions involved in the solar neutrino-producing pp chain.

BS2005(OP) model. The dominant contribution to the solar neutrino flux comes from the pp reaction itself. However, these neutrinos are of very low energies, in the range $0.1 - 0.4\,\mathrm{MeV}$, and, so, are below the sensitivity of most solar neutrino experiments. The other reactions have lower fluxes but produce neutrinos at higher energies, which are therefore easier to detect. In particular, the ${}^8\mathrm{B}$ neutrinos provide a good source for experimental observation.

Solar neutrinos have been utilised by many experiments to explore the wide ranging field of neutrino physics. The neutrinos are created exclusively in the electron flavour state, such that any observation of non-electron flavour neutrinos from the Sun must be a result of flavour oscillation. In addition, the fluxes are very high, as is required for any measurement of neutrinos due to their weak level of interactions. The neutrino energies span a range of more than two orders of magnitude, from < 0.1 to $> 15\,\mathrm{MeV}$. As such, solar neutrino experiments have provided a wealth of information about these fundamental particles.
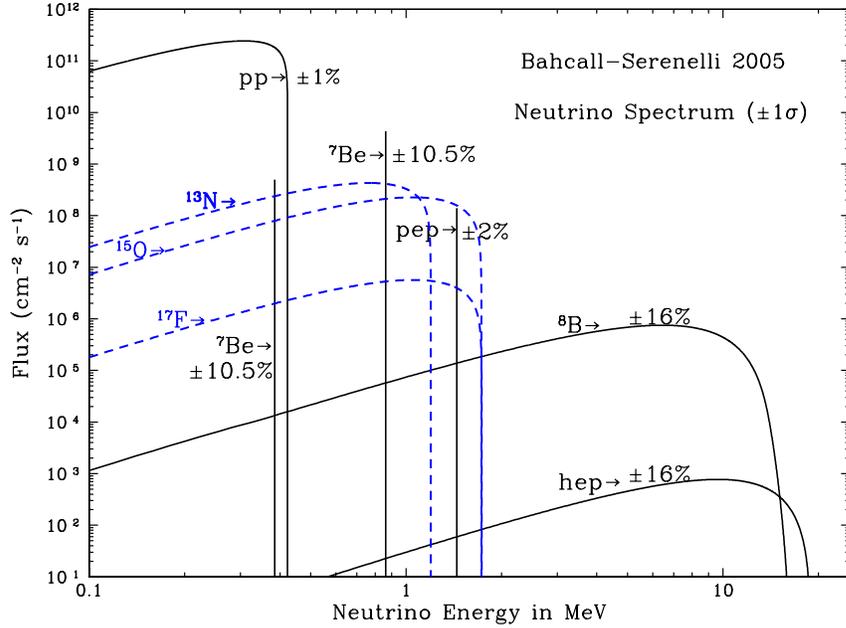
Figure 1.2: The energy spectra of solar neutrinos predicted by the BS05(OP) standard solar model. pp chain neutrinos are represented by solid black lines and CNO neutrinos by dashed blue lines. Taken from [14].

Interaction	Predicted flux / $\mathrm{cm^{-2}s^{-1}}$
pp	5.99 ($\pm$ 1%)
pep	1.42×10^{-2} ($\pm$ 2%)
hep	7.93×10^{-7} ($\pm$ 16%)
^{7}Be	4.84×10^{-1} ($\pm$ 11%)
^{8}B	5.69×10^{-4} ($\pm$ 16%)

Table 1.1: Solar neutrino fluxes from the thermonuclear reactions involved in the pp chain, as predicted by the BS2005(OP) SSM. Fluxes and theoretical uncertainties, given in parentheses and quoted in %, taken from [13].

1.4 Solar Neutrino Experiments

A number of solar neutrino experiments over the years observed a deficit of ν_e in comparison to the flux predicted by the SSM. This phenomenon became known as the 'solar neutrino problem'. The first generation of solar neutrino experiments that observed this problem are described below.

1.4.1 The First Hints

The solar neutrino problem, being the deficit in the flux of ν_e observed from the Sun in comparison to model predictions, was first observed by Ray Davies in the Homestake experiment in the 1960s [15]. Over the following years, several experiments reported similar observations, although the fractions observed in each experiment were not in agreement. These experiments can be separated into two categories according to the method employed to detect neutrinos: radiochemical and water Čerenkov detectors.

Radiochemical Experiments

These experiments used a chemical target to look for the appearance of new isotopes due to the interaction of electron neutrinos on the chosen target nucleus. The neutrino flux was calculated by counting the occurrences of those isotopes over a set period of time and so was an integrated flux above the energy threshold of the reaction.

These experiments were sensitive to electron neutrinos only and provided no information about the energy, direction or time of neutrino detection. However, they were sensitive down to very low energies, of less than $1\,\text{MeV}$. The target consisted of either chlorine or gallium, detecting neutrinos via the following reactions:

$$\nu_e + {}^{37}\text{Cl} \rightarrow e^- + {}^{37}\text{Ar}$$
$$\nu_e + {}^{71}\text{Ga} \rightarrow e^- + {}^{71}\text{Ge}$$
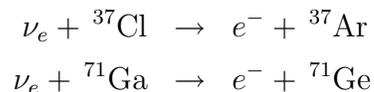
The Chlorine experiment, pioneered by Ray Davies and located in the Homestake mine in South Dakota, famously heralded the first indication of an electron neutrino deficit in the flux from the Sun. With a threshold of $0.814\,\text{MeV}$, this experiment was sensitive to *pep*, *hep*, ^{7}Be and ^{8}B solar neutrinos. The flux was calculated over an exposure time of 1–3 months by flushing the argon out of solution using helium gas and counting the Auger electrons resulting from ^{37}Ar decay. The flux observed was roughly one third of the prediction from the SSM [16].

The gallium experiments detected solar neutrinos from all 5 pp-chain neutrino-producing reactions, with a threshold of $0.233\,\text{MeV}$. The resulting germanium atoms were counted by detecting their radioactive decays. GALLEX [17], later upgraded to GNO [18], and SAGE [19] both observed a deficit of ν_e of roughly a factor of two.

The results from these experiments are summarised in table 1.2 along with the results from the water Čerenkov detectors, which are described below.

Water Čerenkov Experiments

The Kamiokande water Čerenkov experiment [20] detected neutrinos via their elastic scattering (ES) reactions:

$$\nu_x + e^- \rightarrow \nu_x + e^-$$

where x is the neutrino flavour: e, μ or τ. The recoil electron generates Čerenkov light in the H_2O, which is detected using a set of photomultiplier tubes (PMTs) directed at the target volume.

The ES interaction can proceed via both W and Z boson exchange for ν_e, whereas ν_μ and ν_τ can only be detected via Z boson exchange; the interaction is therefore much more sensitive to electron neutrinos, with a reduced sensitivity to the other flavours. The effective energy threshold for water Čerenkov detectors is much higher than the radiochemical experiments, at the order of $5\,\mathrm{MeV}$, but neutrino detection is in real-time, with directional, spectral and time information determined on an individual event basis.

The SuperK experiment was the successor to Kamiokande, employing the same neutrino detection methods but with a much larger fiducial mass of 22.5 ktonnes, in comparison to the 680 tonnes in the original experiment. Both experiments observed a reduced solar ν_e flux in comparison to predictions from the SSM. In addition, SuperK performed a spectral measurement and found the neutrino spectrum to be consistent with an undistorted ^{8}B spectrum. A day-night measurement was also performed, yielding a result consistent with no day-night asymmetry in the flux.

The measured fluxes from these experiments, in comparison to the SSM prediction, are summarised in table 1.2. The ES flux measured by the SuperK experiment [21] was:

$$\phi_{ES}^{SK} = 2.38 \pm 0.05\,(\text{stat})\,^{+0.16}_{+0.15}\,(\text{syst}) \times 10^6\,\mathrm{cm^{-2}s^{-1}} \tag{1.19}$$

Formulating the Solar Neutrino Problem

The fluxes measured in the first generation of solar neutrino experiments are summarised in table 1.2. All the experiments consistently observed a deficit of ν_e in comparison to theoretical predictions, although the exact values measured were not in agreement. The observation of the deficit using three independent detection techniques warred with the success of the SSM in predicting the Sun's behaviour up until that point. The solution that presented itself as the most likely was a fundamental flaw in the understanding of the behaviour of the neutrinos themselves.

The phenomenon of neutrino oscillation was the preferred explanation put forward to explain the observations. The radiochemical experiments detected only ν_e whereas the

Experiment	Detection method	Flux (observed / predicted)
Homestake	Chlorine	0.34 ± 0.06
GALLEX & GNO	Gallium	0.58 ± 0.07
SAGE	Gallium	0.59 ± 0.07
Kamiokande	Water Čerenkov	0.55 ± 0.13
SuperK	Water Čerenkov	0.45 ± 0.08

Table 1.2: Solar neutrino fluxes from the first generation of radiochemical solar neutrino experiments and the major water Čerenkov experiments (pre-SNO), in comparison to the prediction from the BP2000 SSM. Values taken from [22].

water Čerenkov experiments also had a limited sensitivity to other flavours. This differing sensitivity to the three flavours of neutrino could partly explain the discrepancy between the individual experimental results.

However, the results were not consistent with a single set of vacuum oscillation mixing parameters (section 1.1.1). The maximum suppression of the solar ν_e flux in this scenario is 0.5 (equation 1.8), whereas the Homestake experiment in particular observed a significantly greater suppression. In addition, the exact suppression observed appeared to be dependent on the energy range to which the experiment was sensitive.

The possibility of matter effects suppressing the ν_e flux in an energy-dependent fashion, as described in section 1.1.2, seemed the best chance for the phenomenon of neutrino oscillation to offer a resolution to the problem.

1.4.2 The Sudbury Neutrino Observatory

The Sudbury Neutrino Observatory (SNO) was designed specifically to resolve the solar neutrino problem by making the first observation of non-electron flavour neutrino appearance [23]. Utilising the same ES interaction as earlier water Čerenkov detectors, SNO's unique use of heavy water (D_2O) as a target medium allowed it to also detect solar neutrinos via Neutral Current (NC) and Charged Current (CC) interactions on the deuterium nuclei. The ES interaction is sensitive to all neutrino flavours but with enhanced sensitivity to ν_e, by approximately a factor of 6.5. At solar neutrino energies the CC interaction is only sensitive to ν_e, whereas the NC interaction is sensitive to all three neutrino flavours equally. SNO could therefore make a measurement of the integral solar neutrino flux independent of its flavour composition and also of the pure ν_e flux.

The SNO experiment took data in three phases, distinguished by the method used to detect the neutrons from the NC reactions: pure D_2O, D_2O loaded with two tonnes of NaCl and a set of ^{4}He proportional counters. The specifics of the three phases and further details about the SNO detector are described in chapter 2.

The results from SNO were an important milestone in neutrino physics. The publication of the CC flux was the first indication that the solar neutrino flux contained a significant non-electron admixture [24]. Differing from the SuperK ES flux by more than $3\,\sigma$, this was the smoking gun that implied that the ES flux must contain some component of ν_μ and ν_τ. Further results followed: SNO's measurement of the ES flux was in excellent agreement with SuperK and confirmation came with the NC measurement [25], which indicated a very significantly higher ($\times\,2$–3) solar neutrino flux than either the CC or ES measurements, entirely consistent with that predicted by the SSM (as given in table 1.1). Table 1.3 summarises the flux results from the first two phases of SNO.

SNO also performed a measurement of the neutrino fluxes free from any assumptions about the shape of the underlying ν_e energy spectrum [12]. In this analysis, the shape of the neutrino energy spectrum was allowed to vary as a free parameter in the extraction. This result is also given in table 1.3 and is referred to as 'unconstrained'.

	CC	ES	NC
D_2O constrained	$1.76\,^{+0.06\ +0.09}_{-0.05\ -0.09}$	$2.39\,^{+0.24\ +0.12}_{-0.23\ -0.12}$	$5.09\,^{+0.44\ +0.46}_{-0.43\ -0.43}$
Salt constrained	$1.72\,^{+0.05\ +0.11}_{-0.05\ -0.11}$	$2.34\,^{+0.23\ +0.15}_{-0.23\ -0.14}$	$4.81\,^{+0.19\ +0.28}_{-0.19\ -0.27}$
Salt unconstrained	$1.68\,^{+0.06\ +0.08}_{-0.06\ -0.09}$	$2.35\,^{+0.22\ +0.15}_{-0.22\ -0.15}$	$4.94\,^{+0.21\ +0.38}_{-0.21\ -0.34}$

Table 1.3: Summary of flux measurements from the first two phases of SNO: the D_2O and salt phases. All units are $\times 10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The quoted uncertainties are statistical and systematic respectively. Results are taken from [12].

The ratio of the CC to NC fluxes from the unconstrained fit in the second phase of the experiment was:

$$\frac{\phi^{SNO}_{CC}}{\phi^{SNO}_{NC}} = 0.340 \pm 0.023\,(\text{stat})\,^{+0.029}_{-0.031}\,(\text{syst})\,\times 10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1} \tag{1.20}$$

which is significantly less than 0.5, in agreement with the addition of the MSW effect to the vacuum oscillation model. The results from SNO strongly favoured the explanation of solar neutrino oscillation, with the ν_e generated inside the core of the Sun undergoing flavour transformation as they propagate through the Sun's matter and the vacuum of space, en route to the Earth.

SNO's measurement of the energy spectrum from the unconstrained fit was consistent with an undistorted ^{8}B spectrum, as well as with the slightly distorted spectrum predicted by the MSW effect (section 1.1.3). In addition, a day-night analysis was performed and the result was consistent with no asymmetry between the neutrino fluxes measured during the day and during the night. The predicted effects are small and greater experimental accuracy would be required in order to probe them further.

1.4.3 The KamLAND Experiment

KamLAND is a long baseline reactor experiment, detecting antineutrinos from 53 nuclear power plants in Japan and South Korea, via the inverse beta decay reaction in a liquid scintillator target [26]. The mean neutrino energy is approximately 4 MeV and the mean distance of the detector from the reactors is ~ 180 km. As a result, KamLAND is sensitive to the $[\theta_{12}, \Delta m_{12}^2]$ sector of neutrino oscillation, but without complications from the MSW effect.

Using an energy threshold of 3.4 MeV, KamLAND found the ratio of detected events to that expected, assuming no oscillation, to be:

$$\frac{\text{Detected}}{\text{Expected}} = 0.658 \pm 0.044\,(\text{stat}) \pm 0.047\,(\text{syst}) \tag{1.21}$$

which allowed it to make a measurement of the relevant mixing parameters.

The results from KamLAND were in very good agreement with the values of the mixing parameters measured by SNO and other solar neutrino experiments. The combined results from all these experiments place very tight limits on the parameters, as described in the following section.

1.4.4 The LMA Solution

The combination of the observations from the first generation of solar neutrino experiments resulted in four allowed regions of parameter space in the $[\theta_{12}, \Delta m_{12}^2]$ plane. The inclusion of the result from SNO limited this to one region in the upper right corner of the plane, known as the Large Mixing Angle (LMA) solution. The mixing in this region is large but maximal mixing is ruled out with very high significance.

The LMA region corresponds to the adiabatic condition in the Sun, in which the change in density as the neutrino propagates is slow enough that transitions between the matter eigenstates can be neglected. The composition of the instantaneous matter eigenstates is dependent on the changing density of the medium but the admixture of those matter states is determined by the density at the production point, in the Sun's core. To a good approximation, neutrinos are produced in the ν_{2m} state, as described in section 1.1.2.

The results from KamLAND further confirmed the LMA region of parameter space, assuming CPT conservation and the two neutrino approximation. Since matter effects are negligible in the KamLAND experiment, the agreement of the results with those from solar neutrino experiments (which take the MSW effect into account when calculating the oscillation parameters) is strong evidence for the suppression of the ν_e survival probability by the effect of the Sun's matter.

The allowed regions of parameter space are illustrated in figure 1.3, which shows the results from SNO and other solar neutrino experiments before and after the inclusion of the KamLAND results.
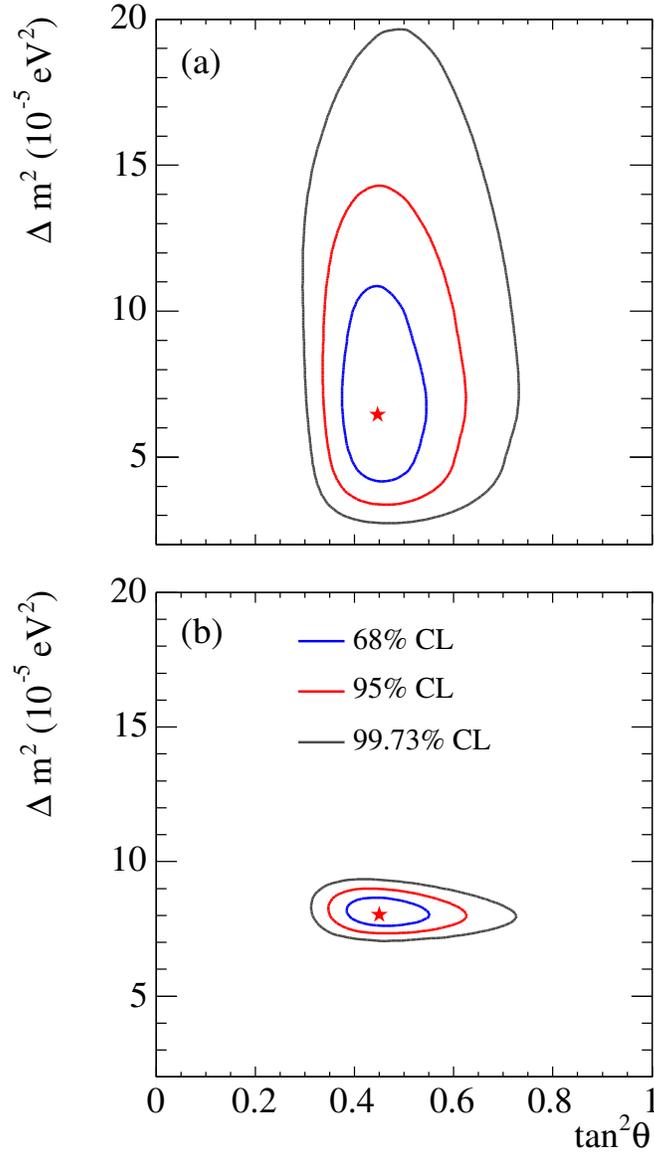
Figure 1.3: The allowed regions of parameter space in the solar neutrino sector of neutrino oscillations, for (a) SNO and solar neutrino experiments only and (b) the inclusion of the KamLAND results. θ and Δm^2 refer to θ_{12} and Δm^2_{12} and C.L. refers to the Confidence Level. Figures taken from [12].

1.5 Lowering the Threshold

The LMA solution has been firmly established as describing the solar neutrino oscillation parameters, θ_{12} and Δm_{12}^2. This solution incorporates the flavour transformation that occurs in the Sun due to the MSW effect, resulting in an energy-dependence to the ν_e survival probability. Figure 1.4 illustrates this dependence for the full three neutrino scenario; in the two neutrino approximation, θ_{13} is small and so $\cos\theta_{13}$ is taken to be 1.0.
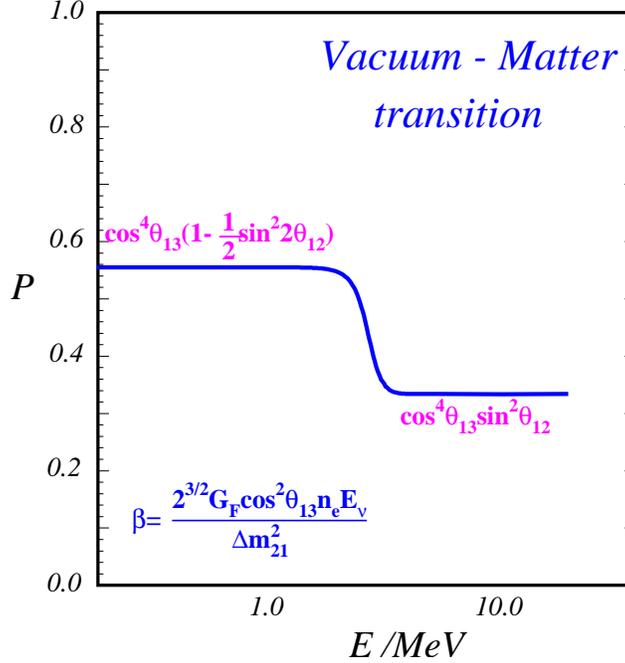
Figure 1.4: The electron neutrino survival probability as a function of energy for the LMA solution. $\beta = \frac{L_V}{L_e}$, where L_V and L_e are the vacuum and matter oscillation lengths as defined in equations 1.6 and 1.10. The values on the x-axis are a guide only. Figure taken from [14].

For neutrino energies below $2\,\text{MeV}$, matter effects are negligible and the survival probability is given by the vacuum oscillation result from equation 1.8. For energies above $10\,\text{MeV}$, the MSW effect dominates and the ν_e survival probability is suppressed by adiabatic flavour conversion in the Sun. In the intermediate region, there is a transition between vacuum and matter dominated effects in which the oscillation term is significant. In this region, a small distortion is predicted in the ${}^8\text{B}$ neutrino spectrum. An experiment with a low enough energy threshold and a high degree of accuracy could probe this effect.

In addition, this transition region is particularly sensitive to new physics, including non-standard interactions and mass-varying neutrinos.

1.5.1   The Low Energy Regime

Although the radiochemical experiments described in section 1.4.1 had very low thresholds for solar neutrino detection, the measured flux was integrated above that threshold and so incorporated the higher energy neutrinos, which are affected by the MSW effect. The Borexino experiment aims to measure the flux of ^{7}Be neutrinos, which are emitted in a monochromatic line at 863 keV. Neutrino detection is via the ES reaction in a 100 tonne organic scintillator target. ^{7}Be neutrinos are the second dominant component of the overall flux from the Sun and are predicted with significantly higher accuracy than the ^{8}B flux, allowing for a high precision test of the low energy regime of the LMA solution. First results from the experiment gave a flux of 47 ± 7 (stat) ± 12 (syst) counts /(day*100 ton), consistent with the LMA-predicted suppression of solar ν_e [27].

1.5.2   Probing the MSW Transition

The water Čerenkov experiments described above probe the higher energy regions of the LMA solution, above 5 MeV. The predicted distortion of the ^{8}B spectrum in the transition region is small and requires a high degree of experimental accuracy for an observation to be made. The distortion increases at lower energies, so lowering the threshold of analysis in these experiments increases the possibility of observing the effect.

As well as searching for the predicted 'MSW-rise' in the neutrino energy spectrum, lowering the energy threshold of the water Čerenkov experiments increases the chance of observing new physics, since the effects become more pronounced at lower energies.

At the Threshold of New Physics

Non-standard interactions (NSIs) such as flavour-changing NC processes could affect the MSW predictions [28]. In the Standard Model, all NC processes conserve particle flavour. The existence of flavour-changing processes could change the MSW resonance condition and, hence, would affect the energy dependence of the ν_e survival probability. The effect of these NSIs tends to shift the predicted rise in the spectrum to lower energies, although some models also cause the rise to steepen, depending on the size of the additional term in the Hamiltonian for the system. This effect becomes more significant at low energies. The current threshold of the SNO experiment is at the upper end of any predicted distortion. A reduced detector threshold and improved experimental accuracy could enable SNO to rule out some of the NSI models.

The scale of the neutrino mass-squared differences is on the order of $0.01 \, \text{eV}^2$. In one of nature's interesting coincidences, this is very similar to the characteristic scale of the apparent 'dark energy' that is causing the accelerated expansion of the Universe

$(0.002 \, \text{eV}^4)$. This has led to the proposal of mass-varying neutrinos (MaVaNs) as a means of linking the two phenomena [29]. The premise is that the summed density of neutrinos and dark energy is constant over time. A coupling is postulated between a sterile neutrino and a light scalar field, called an acceleron. The effective potential of the acceleron is proportional to neutrino density and, as a result, the neutrino mass is dependent on the local neutrino density. Variations in neutrino mass with time could then explain currently unresolved properties of the dark energy, such as its scale relative to that of dark matter. Solar neutrino oscillations offer a possible test of this theory, due to the high density of neutrinos produced in the core of the Sun. The effect of the inclusion of MaVaNs on the energy-dependence of the LMA solution is illustrated in figure 1.5. The ν_e survival probability can be dramatically changed from the LMA-prediction and still be consistent with current measurements. Improved experimental accuracy and a lowered threshold of analysis would allow for further tests of this theory.
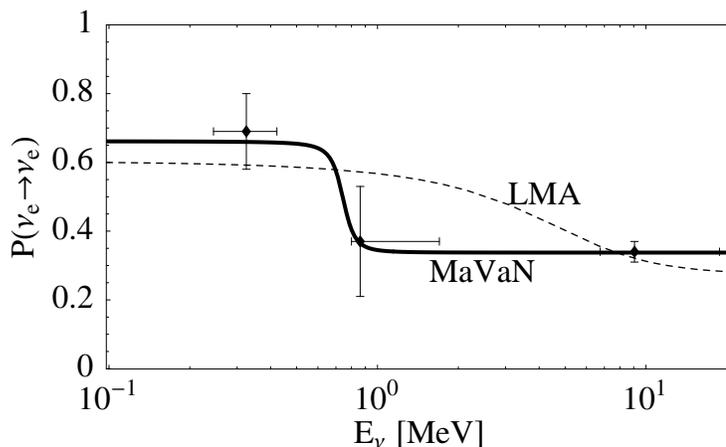
Figure 1.5: The electron neutrino survival probability as a function of energy for the LMA solution and with the additional effect of MaVaNs. Figure taken from [29].

SNO's Low Threshold Analysis

The SNO experiment has already been instrumental in constraining the allowed $[\theta_{12}, \Delta m_{12}^2]$ space, as discussed in the previous sections. The best-fit LMA solution is in good agreement with the results from a number of independent experiments and vacuum oscillation combined with matter effects has been established at a $5.3 \, \sigma$ significance. Although the MSW effect has been a resounding success in terms of predicting the behaviour of neutrinos from the Sun, several other theories, described above, are consistent with the observations. In addition, matter effects in the Sun lead to two predictions, spectral distortion and a day-night asymmetry, that remain, as yet, unobserved.

The aim of the work presented in the following chapters of this thesis is to improve on SNO's previous measurements of the solar neutrino flux. This will be attempted by combining data from the first two phases of the experiment, improving the accuracy of every aspect of the detector model and reducing the energy threshold of the analysis. The resulting increased sensitivity to possible spectral distortions could either confirm or rule out a number of models, from MSW effects to more exotic physics such as NSIs and MaVaN models.

SNO's unique ability to detect ν_e via the CC process allows for a precise measurement of the underlying neutrino spectrum. Although the ES interaction preserves a small amount of information about the incident neutrino energy, the recoil electron from the CC interaction provides a much greater handle on this spectrum. Consequently, the detected CC spectrum is much more sensitive to distortions in the ^{8}B neutrino spectrum.

Figure 1.6 shows the predicted MSW distortion in the CC and ES recoil electron spectra at the best-fit LMA point, in comparison to the undistorted spectra. The distortion causes a rise at the low end of the spectra, with the magnitude of the effect increasing at lower energies. A low energy threshold for analysis is therefore a critical component in searching for possible distortions. The distortion of the CC flux at the best-fit LMA point can be seen to be greater than 8% at 3.5 MeV in the recoil electron energy spectrum.

In order to demonstrate SNO's ability to discern this feature, a Monte Carlo simulation has been used to produce the reconstructed spectra for distorted and undistorted events. The ratio of these are also shown in figure 1.6, for comparison. The finite detector resolution effectively smears out the distortion, reducing the size of the effect. However, at a 3.5 MeV threshold, sufficiently high experimental accuracy would allow for an observation of such a distortion.

1.6 Summary

The observation of neutrino flavour transformation requires the introduction of a neutrino mass term to the Standard Model of particle physics. The understanding of these fundamental particles has progressed far over recent decades, but several unanswered questions remain. Although the phenomenon of neutrino oscillation coupled with matter effects in the Sun has explained the experimental observations with great success, the predicted MSW distortion in the spectrum at low energies has not yet been able to be observed.

This thesis describes an improved spectral analysis of data from the SNO experiment, utilising a lowered energy threshold and an improved understanding and modelling of the detector in order to reduce both statistical and systematic uncertainties. The resulting sensitivity should yield the most precise measurement of the solar neutrino flux yet made
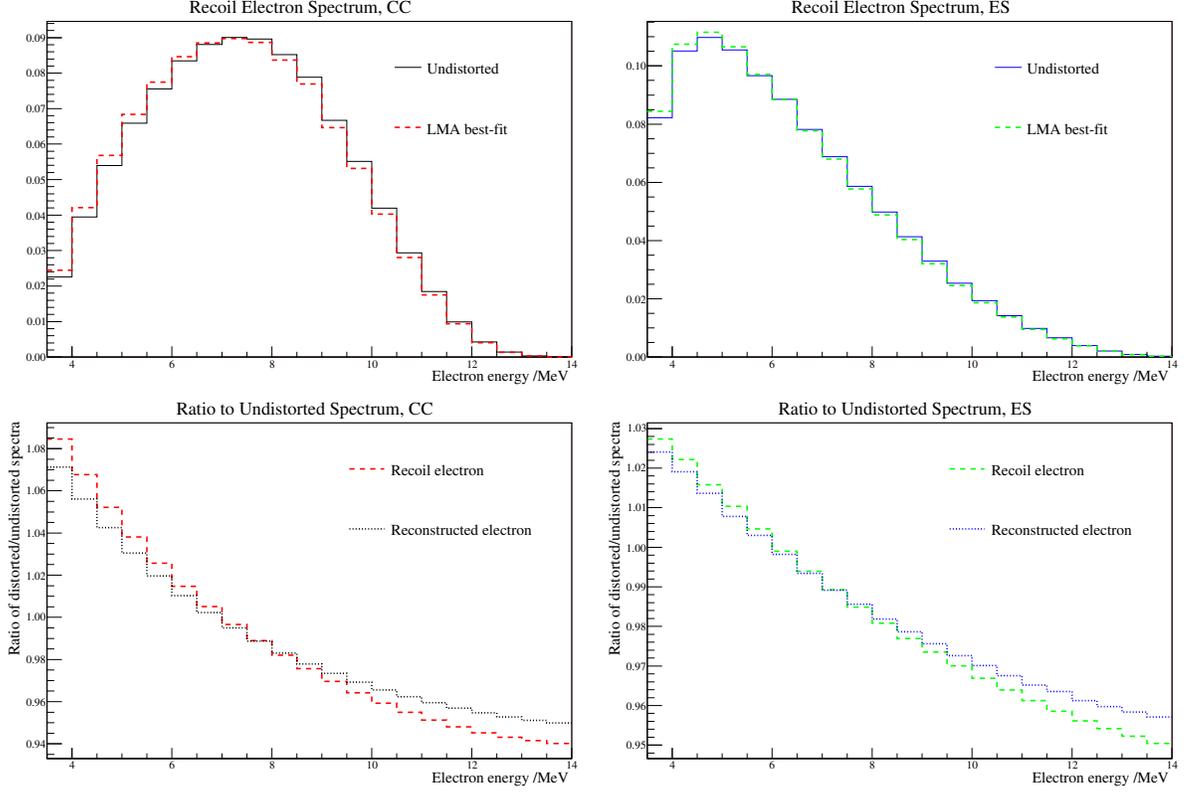
Figure 1.6: Predicted MSW distortion of the CC (left) and ES (right) recoil electron energy spectra. The LMA best-fit point refers to values of $\theta_{12} = 0.45$ and $\Delta m_{12}^2 = 8.0 \times 10^{-5}\,\text{eV}^2$. The upper figures are normalised by area. The lower figures show the ratio of the distorted to undistorted spectra for both the recoil electron energy and the reconstructed electron energy.

and, as such, the best measurement of the neutrino mixing parameters. In addition, the lowered analysis threshold opens the door to searches for possible distortions in the neutrino spectrum. A sufficiently low threshold and high level of experimental accuracy would allow for a test of the predicted MSW-rise of the survival probability at low energies.

In order to meet these goals, the detector must be understood in detail and accurately modelled in the simulation. Each source of uncertainty must also be measured with very high precision. There are essentially three steps to achieving the desired accuracy in the final measurement of the neutrino fluxes and spectrum:

- Improvements in the simulation of the detector

- A more accurate assessment of systematic uncertainties

- An improved extraction of the neutrino fluxes, including a full propagation of all

associated uncertainties

Each of these steps plays a vital role and contributes to the accuracy with which the final neutrino fluxes can be measured. The work presented in this thesis addresses each of these aspects of the analysis in order to make a final measurement of the neutrino fluxes and spectrum.

The introductory chapters describe the relevant background for the work presented here. A detailed description of the SNO detector is given in chapter 2 and chapter 3 describes the motivations behind lowering the energy threshold and the goals of the analysis. The key elements of the analysis are described, along with the improved tools and methods developed in order to meet the challenges of the lowered threshold.

Chapter 4 tackles the first of the three steps by focusing on improving the modelling of the SNO detector in the Monte Carlo simulation. In particular, the modelling of the optical response of the photomultiplier tubes (PMTs) is optimised to better reproduce *in-situ* calibration data. This has a direct impact on the accuracy with which the energy scale can be measured and, thus, the accuracy with which the neutrino spectrum can be extracted.

Chapters 5 and 6 tackle step two, describing in-depth studies of certain systematic uncertainties relating to the measurement of the neutrino fluxes. New techniques are developed to allow the determination of these uncertainties with greater accuracy, with a direct impact on the resulting accuracy in the measurement of the neutrino fluxes.

Chapters 7 and 8 describe the development and verification of the tools and methods required to extract the neutrino fluxes and spectra from the data set in order to complete the final step. The results from the earlier chapters are incorporated into the analysis in the propagation of systematic uncertainties, with a direct impact on the accuracy of the final measurement. The results of an extraction of the neutrino fluxes from the SNO data using the lowered energy threshold of 3.5 MeV in reconstructed electron kinetic energy are reported. The final conclusions of this thesis are then presented in chapter 9.

Chapter 2

The Sudbury Neutrino Observatory

The Sudbury Neutrino Observatory (SNO) experiment consisted of a one kilotonne water Čerenkov detector, buried deep underground in the CRVD-INCO Creighton mine near Sudbury, a small mining town in Canada. SNO's unique use of heavy water as a detection medium allowed it to probe the flux of all active flavours of neutrino from the Sun, thus directly testing the hypothesis that solar ν_e can oscillate to other flavours. This chapter describes the main features of the SNO detector and their relevance to performing a measurement of the solar neutrino flux.

2.1 Interactions in SNO

The SNO experiment was separated into three phases of data taking. These were differentiated by the medium used to detect the neutrons produced by neutrino interactions in the detector.

I. The original target mass was a kilotonne of pure heavy water (D_2O), allowing a measurement of both the total solar neutrino flux and the ν_e component of that flux, as described in the following section. This first phase ran from November 1999 to June 2001.

II. The second phase involved the addition of two tonnes of sodium chloride to the detector, enhancing the neutron capture efficiency from the 30% seen in phase I to nearly 83%. The second phase ended in September 2003.

III. Phase III of the experiment involved deploying 40 ^{3}He proportional counters, or neutral current detectors (NCDs), to provide a method of detecting neutrons that was independent of the results from the first two phases.

The analysis presented in this thesis used data from phases I and II only and so these are the focus of the following sections.

2.1.1 Neutrino Interactions

SNO's target mass consisted of heavy water, which gave the experiment a unique sensitivity to all active flavours of neutrino. The incident neutrinos were detected via three interactions: Neutral Current (NC), Charged Current (CC) and Elastic Scattering (ES).

The NC interaction is equally sensitive to all active neutrino flavours, since it is mediated by the neutral Z-boson, thus allowing SNO to make a measurement of the total flux of solar neutrinos. An incident neutrino breaks up the deuterium nucleus (d) into a free proton and neutron, as shown below and illustrated in figure 2.1 (a). The energy threshold of this reaction is the binding energy of the deuteron, 2.226 MeV.

$$\text{NC:} \quad \nu_x + d \rightarrow n + p + \nu_x \tag{2.1}$$

The CC reaction involves the exchange of a W-boson in the nucleus of a deuterium atom, converting a down quark into an up quark and hence a neutron into a proton. The energy threshold for this conversion is 1.44 MeV. In order to conserve lepton number, the interaction of each flavour of neutrino results in the production of its charged lepton partner (e, μ or τ), as illustrated in figure 2.1 (b). The solar neutrinos detected by SNO are not energetic enough to produce either a μ or τ and so this reaction is only sensitive to ν_e.

$$\text{CC:} \quad \nu_e + d \rightarrow 2p + e^- \tag{2.2}$$

The third interaction by which SNO detected neutrinos was Elastic Scattering (ES), as illustrated in figure 2.2. This is sensitive to all active flavours of neutrino, but with enhanced sensitivity to ν_e, a factor of roughly 6.5 at the energies detected by SNO, since it can occur by exchange of either a W or a Z boson.

$$\text{ES:} \quad \nu_e + e^- \rightarrow \nu_e + e^- \tag{2.3}$$

The CC and ES reactions were observed by detecting the Čerenkov light produced by the recoil electron (see section 2.1.2). In both cases, the energy of this electron is related to the energy of the incident neutrino; in particular, the protons produced in the CC reaction have very low momenta and so the electron carries away the majority of the neutrino energy, minus the Q-value of 1.44 MeV. Therefore, these reactions provided a good measurement of the ^{8}B neutrino energy spectrum. Additionally, the recoil electron from the ES interaction carries information about the direction of the incident neutrino
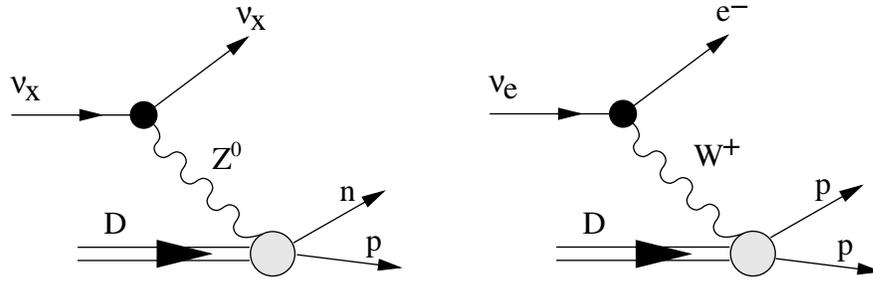
Figure 2.1: Feynman diagrams of the (a) NC and (b) CC neutrino interactions in SNO.
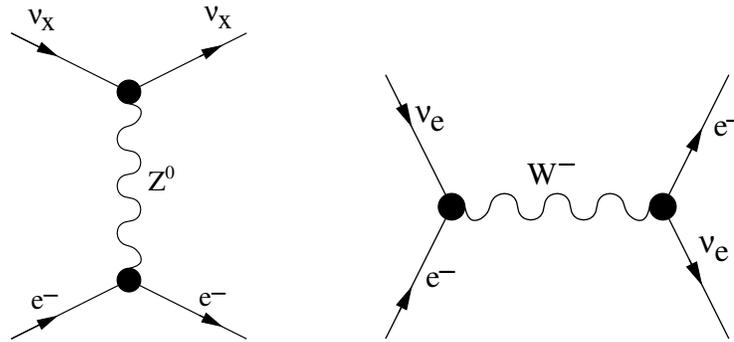
Figure 2.2: Feynman diagrams of the ES neutrino interactions in SNO, by exchange of the (a) Z and (b) W bosons.

and so was used to establish the Sun as the source of the neutrinos, through conservation of momentum. The CC reaction also carries some directional information; this is discussed in more detail in chapter 3.

The signal from the neutron produced in the NC reaction provides no information about either the energy or direction of the incident neutrino. It was detected via a different method in each phase of the SNO experiment. In the first phase, neutron capture on a deuteron produced an excited state of tritium, which decayed with the release of a 6.25 MeV γ. In the second phase of the experiment, in which the target mass was loaded with two tonnes of sodium chloride, the neutron captured predominantly on chlorine. The resultant decay produced a cascade of γs, with a total energy of 8.6 MeV. The γs from these decays were detected primarily through their Compton scatters with electrons, which produced Čerenkov light if the electron was above the threshold (section 2.1.2).

In the third phase of the experiment, a set of 40 ^{3}He NCDs were deployed throughout the detector volume. Neutron capture on the helium produced a tritium nucleus and a free proton and the resulting ionisation was detected by a central anode wire.

2.1.2   Čerenkov Radiation

Neutrino interactions in the first two phases of SNO were observed via their production of Čerenkov light. When a charged particle traverses a medium with a velocity greater than the speed of light in that medium, $v > \frac{c}{n}$, it produces a coherent wavefront of radiation. The minimum energy for which this occurs is known as the Čerenkov threshold.

The CC and ES reactions produced a recoil electron that could generate Čerenkov light directly. In the first two phases of the experiment, the capture of the neutron from the NC reaction produced γs, which could Compton scatter electrons and so produce Čerenkov light if the resulting electron was above the Čerenkov threshold.

The number of Čerenkov photons produced per unit path length, dx, per unit energy, dE, is given by:

$$\frac{d^2N}{dxdE} = \frac{2\pi\alpha}{\lambda^2}\left(1 - \frac{1}{\beta^2 n(\lambda)^2}\right) \tag{2.4}$$

where α is the fine structure constant ($\sim \frac{1}{137}$), λ is the wavelength of the emitted photons, $\beta = \frac{v}{c}$ is the velocity of the particle and n is the refractive index of the medium, which is a function of wavelength and, hence, of photon energy. The number of photons generated at each wavelength is a direct function of the energy of the charged particle. The resulting light is emitted in a cone around the momentum vector of the particle, defined by:

$$\cos\theta_c = \frac{1}{\beta n} \tag{2.5}$$

The refractive index of heavy water is approximately 1.34, resulting in Čerenkov light emitted at an angle of roughly 42° and a Čerenkov threshold of $\sim 767\,\mathrm{keV}$ for electrons.

2.2   The SNO Detector

The SNO detector was located in a large cavity on the 6800 foot (2000 m) level of the CVRD-INCO Creighton Mine in Ontario, Canada. The rock overburden provided shielding from cosmic radiation equivalent to 6010 m of water, reducing the background rate of cosmic ray muons passing through the detector to roughly 70 per day. The one kilotonne target volume of heavy water (D_2O) was housed within a 5.5 cm thick spherical acrylic vessel (AV), with a radius of 6 m. Surrounding the AV was 7400 tonnes of ultra-pure light water, which acted as a buffer region for background radioactivity from the PMTs and the surrounding rock. A cross-section of the detector geometry is shown in figure 2.3.
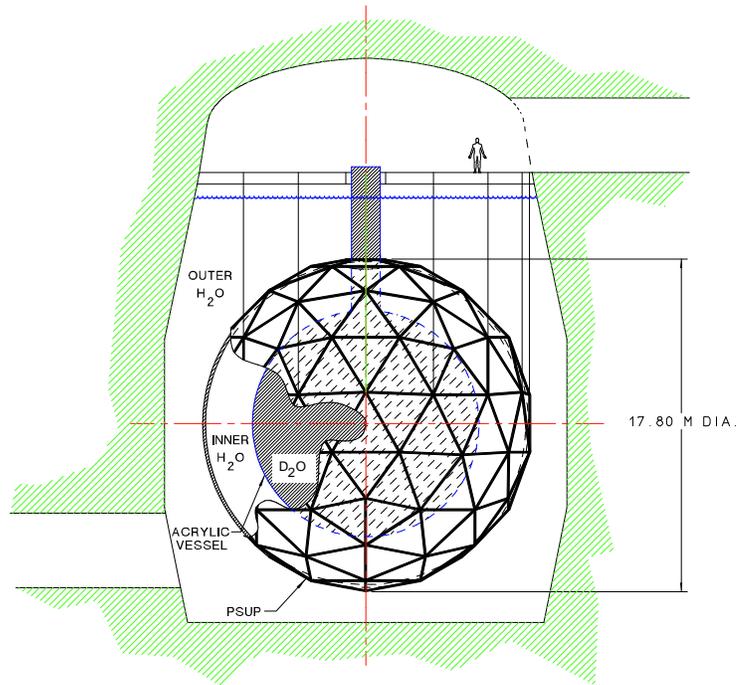
Figure 2.3: Cross-section of the SNO detector geometry. Taken from [30].

Čerenkov light from events within the detector was observed using 9438 photomultiplier tubes (PMTs) that faced the D_2O volume. These were mounted on a $17.8\,\text{m}$ radius spherical PMT support structure (PSUP), which sat concentric with the AV. A further 91 outward looking (OWL) tubes faced the H_2O volume, in order to tag cosmic muon events.

The inner PMTs provided 31% photocathode coverage of the detector. The addition of a $27\,\text{cm}$ diameter light concentrator to each PMT increased the global light collection efficiency to 54%. The response of the PMT and light concentrator arrangement was constant to within 15% out to incident angles of $50°$, after which it dropped sharply, resulting in good coverage of the D_2O volume but low sensitivity to events occurring in the H_2O buffer region.

A $1.5\,\text{m}$ diameter *neck* gave access to the inner volume, allowing calibration sources to be deployed within the detector (see section 2.3.2). For analysis purposes, a Cartesian coordinate system was defined such that the centre of the vessel was at $(x,y,z) = (0,0,0)$ with the neck located symmetrically about the positive z-axis. Further details about the detector can be found in [30].

2.2.1 Electronics

The electronics were responsible for measuring and storing charge and timing information from the PMTs. Each PMT was connected to a separate *channel* in the SNO electronics. Each set of 32 channels was handled by one PMT Interface Card (PMTIC), which received the signals from the PMTs and also provided the high voltage supply. The signals travelled from the PMTIC to the Front End Card (FEC) and 16 cards were grouped together in a *crate*. When a PMT detected light from a Čerenkov event within the detector, the signal travelled down a coaxial cable to the FEC, which processed, digitised and stored the charge and timing information for each PMT. An overview of the electronics is shown in figure 2.4.
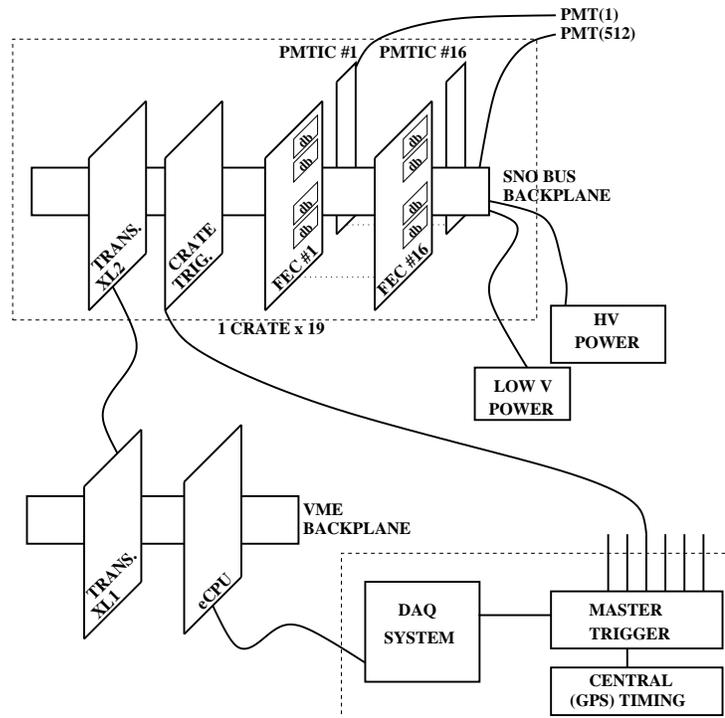
Figure 2.4: Overview of the SNO electronics. Taken from [30].

Individual discriminators determined whether the signal from each PMT was above threshold. When a discriminator fired, multiple types of trigger signals were generated. The SNO trigger system identified events to be read out and stored to disk [31] based on these signals. SNO made use of a variety of trigger types. The main trigger used for physics analyses was the NHIT100 trigger. This identified events based on the number of tubes that fired within a set time window. When a single discriminator fired, a 30 mV square pulse was generated, 93 ns wide and with a 2.5 ns rise time. These were continu-

ously summed over all channels; when 16 hits were detected within the 93 ns window, a global trigger was initiated.

Other triggers utilised in SNO included charge-based triggers, such as the ESumHi trigger. This recorded an analogue copy of the pulse from each PMT, summing the pulses to give the total charge deposited in the detector. The longer rise-time of these pulses relative to the sharp signal from the discriminator made them too slow for analyses of physics events, but this trigger was a useful tool for diagnostics of non-physics backgrounds.

The Pulsed Global Trigger (PGT) was a fixed interval trigger, firing at a rate of 5 Hz. This provided a useful measure of ambient noise in the detector. In addition, external triggers could be activated by the use of calibration sources, providing a way to 'tag' events (see section 2.3.2).

These individual trigger pulses were summed separately over each card and each crate. The crate-summed signals were sent to the analogue master trigger card (MTC/A) to determine the detector-wide trigger sum. When this sum crossed the threshold, the information was passed to the digital master trigger card (MTC/D). If the trigger that crossed threshold had been 'masked in' (meaning the electronics had been set to read out when this trigger fired) then the MTC/D initiated a global trigger (GT). Once a GT was received, the integrated charge and time information for each hit PMT was read out by the data acquisition system (DAQ) and stored on disk.

The time of firing of each PMT was calculated using a time-to-amplitude converter (TAC). The capacitor on the TAC started ramping when a discriminator fired and stopped when a GT was received. This approach is known as 'common stop', in contrast to the 'common start' approach of drift chambers, for example, where the trigger comes first and the timing stops when a hit is observed. If no GT was received after a fixed timeout period, the TAC was reset. This timing information was stored along with the integrated charge, digitised by the Analogue to Digital Converters (ADCs), and a global identification number (GTID), which was used to assemble hits from individual PMTs into event bundles.

The AMB (Analogue Measurement Board) was a particularly useful tool for diagnostics and for identifying instrumental backgrounds (see chapter 6). The AMB received a copy of the summed ESumHi signal and the raw trigger pulse from the MTC/D, which it used to determine at what point in time the signal crossed threshold. It then calculated and digitised the integral of the signal from 70 ns prior to this point until 130 ns after this point, along with the derivative and the peak value. This resulted in a measurement of the total charge deposited in a single event, independent of individual channel properties.

The trigger system has been shown to run at a very high efficiency. Using a threshold of 16 hits within a 100 ns period, the NHIT100 trigger was 100% efficient at total energies

of 4 MeV [32]. Although the SNO detector could operate continuously, for organisational purposes the events were grouped into periods of time known as 'runs', with a maximum duration of 24 hours per run.

2.2.2 Water Systems

A vital consideration in SNO was the reduction of the level of background events. The detector was cooled to a temperature of 10°C, both to reduce the PMT noise rate and to minimise the possibility of biological growth inside the detector.

During construction, all the materials in SNO were selected for radioactive purity. Radioactive decays contributed to the level of background events and to photodisintegration of deuterons within the detector and, so, it was critical to maintain as high a level of purity as possible. The D_2O loaned to SNO by Atomic Energy of Canada was provided with a greater than 99.9% level of isotopic purity. Extensive systems were in place to maintain and monitor the levels of radioactive contamination in both the heavy and light water.

In advance of filling the AV, the heavy water was degassed to remove radon and passed through both ion-exchange columns and a reverse osmosis system to remove impurities. To maintain the level of purity even after the D_2O was introduced into the detector, it was regularly recirculated and passed through ion absorption columns and ultra-filtration units to remove any fine particles that could have been introduced during assay proceedings.

Two radiochemical assay techniques were employed to monitor the levels of radioactive contamination in the water. The first assay technique used a polypropylene column filled with beads coated in a manganese oxide to strip radium and lead contaminants from the water [33]. The radon daughters from the decay of these isotopes could then be counted offline. This method had an extraction efficiency of roughly 95% in the D_2O phase of the experiment and 81% in the salt phase.

The second method involved circulating the water through columns containing microfilters coated with hydrous titanium oxide to remove heavy ions such as thorium, radium and lead [34]. The columns were then eluted with hydrochloric acid and the resulting eluate was concentrated and mixed with a liquid scintillator. The activity was then measured using a set of β-α delayed coincidence counters. The efficiency of this method was strongly affected by the addition of salt to the detector and so its use was predominantly in the D_2O phase.

In addition to these *ex-situ* techniques, *in-situ* measurements of the levels of radioactive contamination in the water were made using PMT Čerenkov data. The results from these analyses showed good agreement with the assay results [35].

The light water in the PSUP region of the detector acted as a buffer against external radioactivity and so it was important to maintain a low level of constituent radioactive contaminants. Before entering the detector, the water was degassed to remove oxygen and nitrogen and passed through a set of filters and ion absorption columns. UV lamps were used to kill off any bacterial growth before cooling the water to 10°C. Removing all gas was shown to lead to PMT breakdown due to low partial pressure in the connectors and so the water was regassed with nitrogen. To maintain the levels of purity once the H_2O was in the detector, it was continually recirculated and assayed, using similar techniques to those employed for the D_2O, and passed through the purification system as described.

A protective layer of nitrogen gas was used as a physical barrier to prevent the radon-rich mine air from contaminating either the H_2O or the D_2O.

2.3 Calibrations

In order to fully understand and interpret the data taken within the SNO detector, a rigorous calibration programme was performed. The direction and trajectory of events within the detector were reconstructed based on the charge and timing of individual PMTs. Calibration of the electronics was therefore vital in order to maintain the accuracy of these calculations. Further detector calibrations were performed in order to understand and evaluate the response of the detector to different event types as a function of both energy and position within the detector volume. This was a critical step in understanding the topology of both signal and background events and therefore affected the ability to identify neutrino events in the data set.

2.3.1 Electronics Calibrations

A photon incident on the PMT photocathode can trigger the release of a photoelectron, which is accelerated by an applied electric field towards the first in a series of dynodes. The incidence of this electron triggers a cascade in the dynode stack, with the exact number of electrons released at each dynode varying according to Poisson statistics, resulting in a distribution of pulse heights for the response to a single photon. The *gain* of individual PMTs within SNO varied by up to 10%, thus affecting the response of the detector.

The electronics calibrations were used to determine the basic charge and timing information of each PMT and the front-end electronics. A calibration of fundamental channel properties such as the discriminator thresholds was performed whenever an electronics component, such as the FEC, was replaced. Two sets of electronics calibrations were performed on a more regular basis, to track time-dependent changes:

- **ECA:** The Electronics Calibration was performed bi-weekly in order to measure the variation of channel properties over time. The DAQ sent a pulse to fire the individual discriminators and then a global trigger was forced at a fixed time delay. This allowed a measurement of the pedestal value of each channel (the "zero charge" reading from the ADCs) and the time slope of the TAC pulse.

- **PCA:** The PMT Calibration was performed monthly to measure the time-response of each PMT. Individual discriminators fired when the leading pulse edge from that PMT crossed a fixed threshold value. The time at which this occurred was dependent on the pulse amplitude: a larger pulse caused the discriminator to fire earlier. Therefore, variations in the pulse height recorded by a single PMT led to an effect called "discriminator walk", in which the PMT firing time could vary by as much as 2 ns. An approximately isotropic laserball source was deployed to measure the relation between the deposited charge and the firing time, resulting in a charge-dependent channel-by-channel correction.

2.3.2   Detector Calibrations

Calibration sources were deployed within the detector in order to measure the response to signal and background event characteristics. Several aspects of the detector response had to be measured, including the global light collection efficiency, the angular response of the PMTs and light concentrators, the optical attenuation lengths, the energy response of the detector as a function of position within the volume and also the acceptance of background events. This required the use of a range of calibration source types. These were deployed through the neck of the AV using a manipulator system, as illustrated in figure 2.5. The pulley system that controlled the position of the source allowed for two modes: single and dual-axis deployment. The former involved deploying the source along the vertical, z-axis of the detector, using ropes fixed at (-16.03, 23.5) cm in the (x,y) plane. This was a very accurate mode of deployment, allowing the source position to be known to within 2 cm. In order to deploy sources off the central axis of the detector, the side ropes were used to move the source in the x-z and y-z planes, where the source position could be determined to within 5 cm for central runs and 10 cm for runs near the edge of the volume, where the most strain was put on the ropes and pulleys of the deployment system. A number of guide tubes were also available for deploying sources outside the AV, in the H_2O volume inside the PSUP.

Both background events and neutrino interactions contaminated the calibration data and so a particularly important feature for calibration sources was the ability to identify events originating from the source. This was achieved in a number of ways, one of which
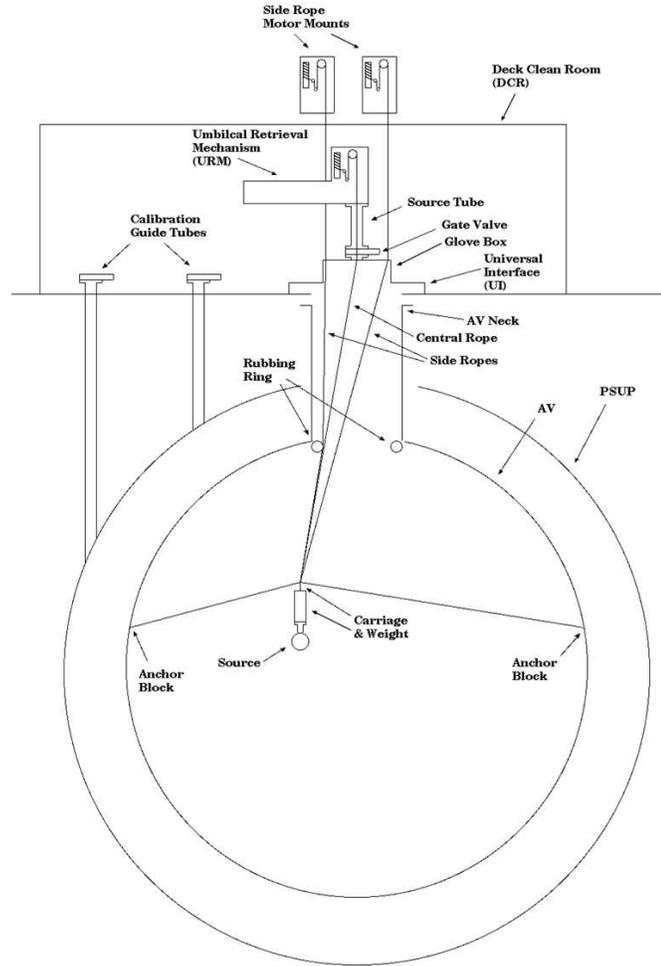
Figure 2.5: Deployment of calibration sources in the SNO detector; taken from [36].

was to use an external trigger to 'tag' source events. The main features of the calibration sources deployed within the SNO detector during the first two phases of data taking are described below.

The Laserball

This triggered, isotropic, multi-wavelength laser source was deployed throughout the detector volume in order to measure optical properties of the detector such as the PMT angular response and the media attenuation lengths. It was also used to measure individual PMT timings during the PCA calibration, as discussed above, and to measure PMT-by-PMT efficiencies.

The ^{16}N Source

The ^{16}N source was a high rate, monoenergetic γ source [37]. The β decay of ^{16}N produced an excited state of ^{16}O, the de-excitation of which generated a 6.13 MeV γ that produced Čerenkov light by Compton scattering an electron. The source was housed in a steel container to absorb the emitted β. A 3 mm thick scintillator sleeve on the inside of the container allowed a small PMT inside the can to trigger on the β, which was then used to tag events from the source, thus producing a clean data sample. This was particularly important for the reconstruction and sacrifice studies presented in chapters 5 and 6.

The ^{16}N source was the primary source used for energy calibration within the detector, since the 6.13 MeV γ tended to Compton Scatter an electron with an energy close to the peak of the CC spectrum (see figure 3.6). As such, this source was deployed in a wide range of positions throughout the detector volume.

Although this was a monoenergetic source, a spectrum of energies was observed for the detected events, partly due to the nature of the Compton scattering involved but also to fluctuations in the energy of the emitted gamma and the finite energy resolution of the detector. This allowed the source to be used for energy-dependent studies across a limited range of energies. However, the high rate also caused event pile-up, particularly at higher energies. This happened when two or more events occurred close enough together in time that the detector could not distinguish between them and they were recorded as one event. To remove the majority of these events, a cut was placed at 10 MeV on all ^{16}N events. Although pile-up was not modelled in the Monte Carlo, this cut was placed on both data and Monte Carlo events, to maintain the comparability of the two.

The ^{8}Li Source

The ^{8}Li was a triggered source that produced a β spectrum with an end-point of $\sim$14 MeV [38]. The energy spectrum produced by ^{8}Li is similar in both shape and end-point to that of the ^{8}B CC signal, since both isotopes decay to the same excited state of ^{8}Be. The source was housed in a steel decay chamber, carefully designed to be clean and robust in order to protect the purity of the D$_2$O, and at the same time to provide a nearly undistorted electron spectrum. The chamber was spherical in design, with a radius of 6.35 cm and a wall thickness of 0.6 mm, corresponding to an electron energy loss of approximately 1 MeV in the range of interest.

Events were tagged by utilising the scintillation properties of the helium gas that filled the chamber. The resulting signal was small and consisted mostly of UV light, which falls at the very low end of the PMT sensitivity range (see figure 4.6). To enhance the signal, the inside of the sphere was painted with a reflective, white, titanium-oxide based paint and coated with a thin layer of tetraphenyl butadiene (TPB), which acted as a

wavelength shifter. The helium also contained a small admixture of nitrogen ($\sim 0.1\%$) to act as an additional wavelength shifter. The daughter nucleus decays promptly, producing two α particles nearly back-to-back with a mean energy of 1.6 MeV each. These particles have a range of about 4 cm in helium gas and so deposited a large proportion of their energy. The resulting scintillation signal was detected with a PMT inside the chamber.

This source was used to confirm results from the ^{16}N studies of the energy response. It was also used to study the detector response differentially with energy. It was particularly important for studies such as the energy-dependence of the detector fiducial volume described in chapter 5.

The ^{252}Cf Source

Approximately 3.1% of ^{252}Cf decays are spontaneous fission, which release a burst of neutrons. This isotope was therefore used as a neutron source, in particular for measurements of the neutron capture efficiency. The source was housed in a plastic can to absorb the βs released by fission fragments. An algorithm was developed, known as 'QBurst', to identify the neutron bursts from the source and to use a timing cut to remove fission γs, which tended to generate light more quickly and therefore form the first event of each burst [39]. Some contamination from such source γs remained in the data, which could be confused with the γs from neutron captures. In addition, neutrons capturing back on the acrylic of the source and the manipulator stem above it produced high energy γs that mimicked neutron events. Care therefore had to be taken when using this source for analyses, such as those discussed in chapter 6.

Acrylic Sources

As described in chapter 3, the dominant source of physics backgrounds in the detector was radioactive decay of isotopes in the ^{238}U and ^{232}Th decay chains, in particular the decays of ^{214}Bi and ^{208}Tl. The full decay schemes for ^{238}U and ^{232}Th are given in appendix A and those for ^{214}Bi and ^{208}Tl are shown in section 3.2.

Sources constructed from ^{226}Ra and ^{232}U were used to model these decays, to calibrate the detector's response to these primary sources of low energy background events. ^{226}Ra is in the ^{238}U decay chain and is therefore referred to as the uranium source. This source was used to generate ^{214}Bi decay events. ^{232}U is in the ^{232}Th decay chain, resulting in ^{208}Tl decay events, so is referred to as the thorium source. Both the uranium and thorium sources were encapsulated in acrylic for deployment within the detector. No method of tagging events existed for either source. The use of these sources to measure background acceptance is discussed in more detail in chapter 7.

The pT Source

The pT source was a tiny accelerator that generated 19.8 MeV γs from the $p+^3\text{H} \rightarrow {}^4\text{He}+\gamma$ reaction. This source was used to measure the response of the detector at high energies. It was not deployed in the salt phase due to the copious number of neutrons that would be produced. This was an untagged source, but cuts on the position and direction of events could be used to reduce contamination in the data set.

The AmBe Source

This source produced a neutron with a 4.14 MeV γ in coincidence. Although it was not a triggered source, events could be tagged by requiring the coincidence, reducing the levels of contamination considerably. This source was used to verify measurements of the neutron capture efficiency in the salt phase; it was not deployed in the D_2O phase.

The Radon Spike

A controlled injection of radon into the D_2O and H_2O volumes allowed a measurement of the detector's response to uniform, isotropic low energy background events such as those from ^{214}Bi contamination in the water. There is no way to directly trigger on a containerless source, but the high rate of events made the contamination from neutrino and other background events a negligible effect.

2.4 Software

Data produced in the SNO detector, both from regular neutrino running and from calibration source deployment, was processed using the SNO Monte Carlo and ANalysis software package, SNOMAN [40]. SNOMAN was used both for data analysis and for Monte Carlo simulations of the SNO detector. Primarily written in FORTRAN 77, SNOMAN makes use of the CERNLIB ZEBRA memory management data structure [41].

The processing of SNO data utilised a full database, SNODB, of all the information necessary to fully describe the detector condition over time, including the electronics and optical calibration constants. SNODB was based on the CERNLIB HEPDB software package [42] and consisted of *titles* banks containing the detector configuration information. Each bank had a validity range so that the correct calibration constants were applied to the PMT data according to when a specific run was taken.

The Monte Carlo simulation is an incredibly complex piece of code, intended to reproduce the detector to such an extent that the simulated events could be processed in an identical fashion to the true data. The simulation models all the relevant interactions in

SNO, including the signal neutrino interactions but also the event types from each of the calibration sources as well as the primary low energy background events from radioactive contaminants and additional sources of background such as atmospheric neutrinos and cosmic muons. SNOMAN uses the EGS4 (Electron-Gamma Shower) code to handle the propagation of electrons and gammas through the detector [43]. Neutron propagation up to energies of 20 MeV is handled by the MCNP neutron transport code developed at the Los Alamos National Laboratory [44].

SNOMAN contains a detailed description of all the significant detector geometry, including the AV, the supporting ropes, the neck, the PSUP and the geometry for the deployment of calibration sources within the detector. In addition, the DAQ, the electronics and the trigger system are all modelled, including effects such as PMT noise and discriminator walk. The algorithms used to reconstruct the position, direction and energy of both data and Monte Carlo events within the detector are also contained in the SNOMAN software.

The optical response of the PMTs is an important aspect of the overall detection efficiency. The detection efficiency for photons as a function of their angle of incidence on the PMT is known as the angular response. A detailed model of this feature was incorporated in the Monte Carlo simulation, propagating the optical photons through a full model of the PMT and light concentrator geometry, simulating both reflections from the concentrator petals and the PMT glass as well as the wavelength dependence of the photocathode response. A full study of the accuracy of this model is presented in chapter 4. Instead of employing the full PMT simulation, the energy reconstruction algorithms used a faster phenomenological model based on laserball measurements, called *'Grey Disk'*. Chapter 4 also presents studies of the agreement between these two models.

2.5   SNO Observables

The charge and timing information from individual PMTs was used to measure the position, direction, energy and isotropy of each event occurring within the SNO detector. Further details on the position and energy reconstruction algorithms used to calculate these quantities are given in chapter 3. The energy reconstruction algorithms made use of the number of tubes that recorded a hit in a fixed time window, the *'NHit'* value, which is proportional to the number of Čerenkov photons produced in an event. There is not a one-to-one mapping between NHit values and the energy of an event, since the number of PMTs that fired varied according to Poisson statistics and the exact position and direction of the event. Since the true energy of an event is not known, throughout this thesis the energy referred to is that produced by these reconstruction algorithms,

which is known as the *effective kinetic energy* (T_{eff}) of an event.

The isotropy of an event is a useful tool for distinguishing different event types. In particular, neutron events in the salt phase resulted in a more isotropic distribution of PMT hits than electron events, since neutron capture on chlorine produced a cascade of γs generated isotropically from the excited nucleus. A number of isotropy parameters have been utilised in SNO analyses. The one used in the analyses presented here, β_{14}, was developed specifically for separating neutron and electron-type events in the salt phase [45]. β_{14} is a linear combination of the average value of the first and fourth Legendre polynomials of the cosine of the angle between each pair of hit PMTs in an event.

An important aspect of the data taken in SNO is whether it displays any variation over time. The time information recorded for each event came from two independent clocks, for redundancy purposes. A 10MHz clock was synchronised with a GPS receiver on the surface and a 50MHz clock was based on a local oscillator. SNO used a universal time scale known as the 'Julian Date'. This is a measure of the time elapsed since 00:00:00 GMT on January 1st 1975, which is offset by 5 hours from the local time in Sudbury. This measure of time is used for all the time-dependent studies presented in the remainder of this thesis.

2.6   Summary

This chapter has described the major features of the SNO detector and how they pertain to a measurement of solar neutrino events. A full understanding is required of each aspect of the detector in order to interpret the recorded data. This is achieved through a rigorous calibration programme, which is used to fine tune the modelling of the detector in the complex Monte Carlo simulation package. The next chapter describes how the SNO data set can be used to measure the ^{8}B neutrino spectrum down to lower energies and with greater accuracy than that achieved in any previous analysis.

Chapter 3

A Spectral Analysis of SNO Data

This chapter describes the motivation behind performing a spectral analysis of the neutrino data collected by SNO. Lowering the energy threshold in order to increase the number of neutrino events within the analysis window raises many new challenges and takes us closer to the limitations imposed by both hardware and physics within the detector. The methods developed to deal with these challenges are discussed and the proposed analysis path is defined.

3.1 Motivation

SNO's measurements of the ^{8}B neutrino flux have already greatly constrained the allowed $[\theta_{12},\ \Delta m_{12}^2]$ parameter space, as discussed in chapter 1. A more precise measurement of the three neutrino signal fluxes would further constrain this space, with the potential to either rule out or support various theories.

One aspect of improving on SNO's current measurements was a reduction of the associated statistical uncertainty by increasing the number of neutrino events in the analysis window. To this end, data from the first two phases of SNO, the pure D_2O and the salt phases, were combined, resulting in a total livetime of over 681 days. In addition, the energy threshold of the analysis was lowered from $5.5\,\mathrm{MeV}$, as used in previous publications, to $3.5\,\mathrm{MeV}$, earning this analysis its acronym of *'LETA'*: the Low Energy Threshold Analysis. As a result, the number of NC events in the analysis window was increased by more than a factor of two over previous analyses of the salt phase data set [12].

A further avenue open for improvements was in the accuracy with which the detector was modelled. Revisiting each aspect of the detector, increasing the understanding of its behaviour and improving the modelling accordingly allowed the systematic uncertainties associated with any measurement to be substantially reduced. In addition, improved analysis techniques allowed a better use of the available information, for example the

improvements in energy reconstruction described in section 3.3.2 significantly reduced the contamination of low energy background events in the analysis window. A more rigorous and correct handling of all the uncertainties, in particular a new method of utilising the likelihood from the signal extraction in the calculation of the contour plot, also contribute to the accuracy of the final measurement.

The physics potential of an improved spectral analysis of SNO data is very exciting. This low energy spectral measurement of the ^{8}B flux has the potential for greater precision than any single phase measurement from the SNO experiment, including that expected for the final NCD phase analysis. It could therefore provide the best constraints on the neutrino oscillation parameters. In addition, lowering the analysis energy threshold increases the sensitivity to spectral distortions predicted by the MSW effect. The CC spectrum is most sensitive to these distortions and, hence, was the focus of this analysis.

3.2   Limitations

Both hardware and physics-based limitations imposed a lower bound on the energy window used for the LETA analysis. An in-depth study of the trigger efficiency at low energies was performed [32], concluding that the triggering system was reliable down to effective electron kinetic energies of 3 MeV. The physical limitations were more complicated to circumvent, requiring the development of new analysis tools and techniques.

The dominant physics background in the SNO data set arises from radioactive contamination. The materials used were specially selected for radiopurity, but naturally occurring radioisotopes such as ^{238}U and ^{232}Th were still present in trace quantities in all components of the detector. As these elements decay, particles can be released that are energetic enough to generate Čerenkov light above the SNO threshold. The resulting background events could mimic both electron and neutron-like neutrino signals. Decays releasing βs above the Čerenkov threshold are almost indistinguishable from low energy CC events. Decay branches consisting of γs above an energy of 2.23 MeV could cause photodisintegration of a deuteron, producing a free neutron that could be confused with those from NC neutrino interactions. In addition, α particles produced in the decay chain could interact with nuclei such as ^{2}H, ^{13}C, ^{17}O and ^{18}O, producing further neutron background events. The cross section for these (α,n) reactions is very small; they occurred in both the D$_2$O and H$_2$O regions but were more common in the acrylic vessel itself.

Due to their relatively high natural abundance, long half-lives and the energy characteristics of their decays, only ^{238}U and ^{232}Th were of concern when considering the SNO data set. The dominant background to neutrino signals in the LETA energy window comes from ^{214}Bi in the ^{238}U chain and ^{208}Tl in the ^{232}Th chain (the full decay chains

are presented in appendix A). The decay chains of both daughters contain low energy β-γ events that can mimic the neutrino signal. The ^{214}Bi and ^{208}Tl decay chains are given in figure 3.1. Over 99% of ^{214}Bi decays produce a β event, which could generate Čerenkov light in the detector and mimic the CC signal. The ^{214}Bi decay scheme has a low probability (approximately 2.5%) of emitting a γ energetic enough to photodisintegrate a deuteron. The ^{208}Tl decay chain contains a 2.614 MeV γ, which could photodisintegrate a deuteron but in fact, due to the energy dependence of the cross sections, was much more likely to undergo Compton scattering.
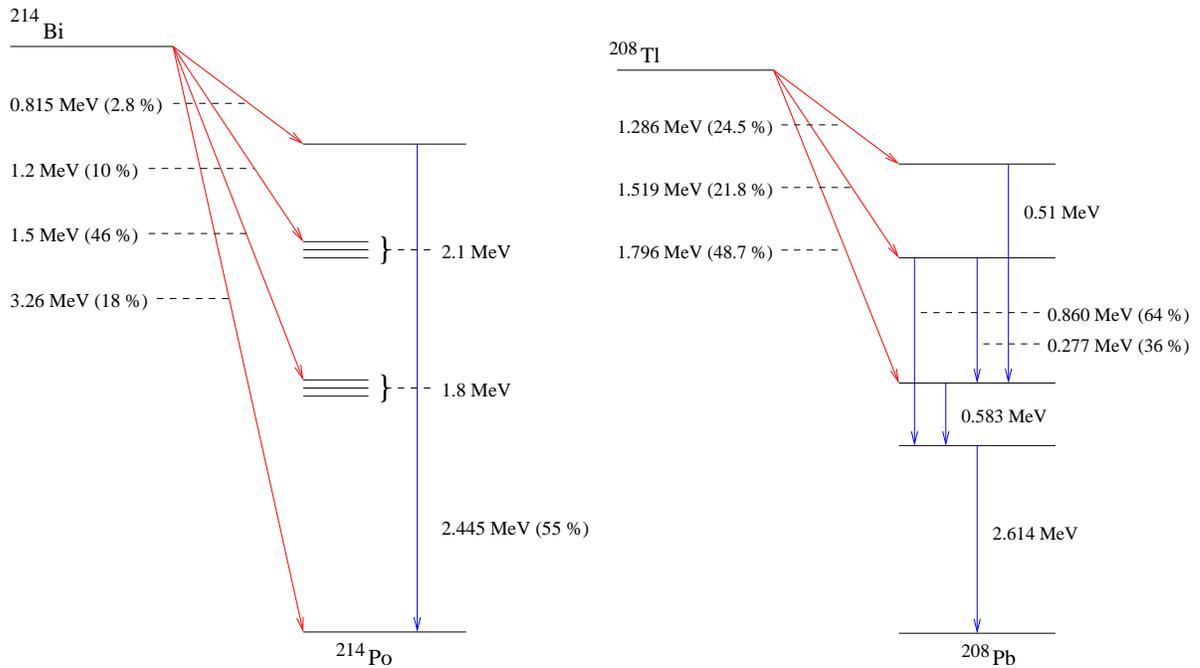
Figure 3.1: The radioactive decay chains for ^{214}Bi (left) and ^{208}Tl (right). Vertical blue lines indicate γ transitions and diagonal red lines indicate βs. The ^{214}Bi scheme has been simplified to show only the relevant transitions. Figures courtesy of H.M. O'Keeffe [46].

Background events were divided into two classes: internal background events being those occurring within the D_2O volume itself and external events being those occurring outside the D_2O, in the AV, the H_2O or the glass of the PMTs. Both the D_2O and H_2O were highly purified. They were circulated regularly and a series of radioassay techniques were employed to monitor the levels of radioactive contamination. A set of filtration systems was in place to reduce the high levels of radioisotopes present in the mine's air and a barrier of nitrogen gas was maintained between the D_2O and the air of the laboratory to reduce radon ingress. Despite all these precautions, some level of contamination remained in the data set. The Čerenkov light from these background events is difficult to distinguish

from genuine neutrino events. Although the energy spectrum of the background events falls off very steeply, as illustrated in figure 3.2, the event rate at the low end of the analysis energy window is still dominated by these events. In addition, a non-zero energy resolution increases the number of events seen in the analysis window. This made the process of energy reconstruction a vital one when considering lowering the analysis energy threshold. This is discussed in more detail in section 3.3.2.
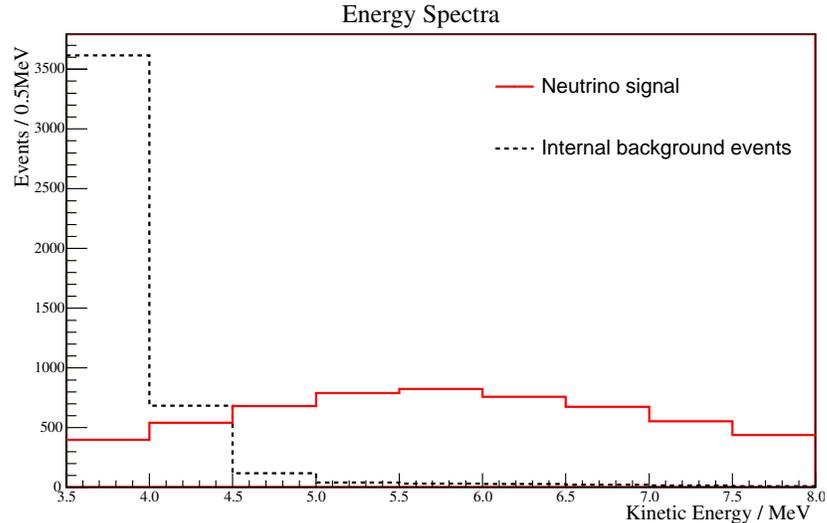
Figure 3.2: The Monte Carlo predicted energy spectra of internal ^{214}Bi and ^{208}Tl events in comparison to that for the neutrino signal (CC + ES + NC). The spectra are normalised to the numbers of events expected in the salt phase data set, using *ex-situ* estimates for the backgrounds and approximate results from previous publications [12] for the signals.

3.3   Towards Lowering the Energy Threshold

In order to meet these challenges, various aspects of the Monte Carlo simulation were improved and a number of new tools and methods were developed for the LETA analysis. These were employed in addition to existing techniques to maximise the separation of signal and background events. In addition, a number of improvements to the data processing procedure were incorporated, described in full in [47], including the addition of a new crosstalk cut to remove a possible source of instrumental background events.

3.3.1   Improving the Model

Many improvements were made to the simulation code to optimise it for the LETA analysis. Some of these were incorporated for previous publications [48] but were not utilised in

the original analyses of the single phase data sets [12],[49]. In particular, many improvements were made to the modelling of the PMTs in the simulation. The timing spectrum was corrected to match bench-top measurements and a new, more realistic model for the charge spectrum was developed based on SNO data. This was then used to apply a threshold cut on individual PMTs in the simulation. Laserball data was used to measure the relative tube-by-tube efficiencies in the detector and this information was included to simulate individual channel efficiencies in the Monte Carlo code. These efficiencies appeared to have a dependence on the z position of the individual PMT, which was not previously modelled in the Monte Carlo simulation, so the inclusion of this effect improved the energy resolution and reduced position-dependent differences between data and Monte Carlo results. New optical constants were also derived using these channel efficiencies.

To bring the amount of prompt and late light in the simulation into agreement with that observed in calibration data, the reflectivity of the concentrators was tuned to increase the number of reflected hits [50]. This allowed both energy estimators (discussed in more detail in section 3.3.2) to use the same global collection efficiency. In addition, the response of the PMT and light concentrator arrangement to photons incident at various angles was tuned to reproduce calibration data. The methods developed for this analysis and its impact on the modelling of the detector are presented in full in chapter 4. This improvement to the simulation was shown to significantly reduce the radial bias in the energy scale uncertainty, to the extent that its contribution to the overall energy scale uncertainty could be neglected in comparison to other effects.

3.3.2   Energy and Position Reconstruction

The LETA analysis made use of several algorithms to reconstruct the energies and positions of events within the detector. The position fitters were utilised for previous SNO analyses as was one of the energy fitters, although improvements were made for the purposes of this analysis. The second energy fitter was developed specifically for LETA. Each of these fitters is summarised below.

Energy Reconstruction

The spectrum of background events from radioactive decays in the SNO detector drops steeply with energy. However, misreconstruction of low energy events to higher energies increases the number of events in the analysis window. Improving the energy resolution of the detector was therefore a vital step in improving the signal to background ratio.

Previous analyses used an energy reconstruction algorithm known as RSP. This algorithm uses only the prompt light from an event to estimate its energy, where a prompt hit is defined as one in which the photon reached the PMT within a window of 10 ns either

side of the predicted time, calculated from the reconstructed position of the event and the average speed of light in the intervening media. The reasoning for this was that late light (that which was reflected or scattered before reaching the PMT) is difficult to model and accounts for only 12% of the total PMT hits. However, a 12% increase in the light collection would lead to an improvement of up to 6% in energy resolution, which would significantly reduce the number of background events contaminating the signal window.

A new energy fitter, FTK, was developed for the purposes of the LETA analysis [51]. As well as taking into account both scattered and reflected light, FTK employs a maximum likelihood technique to estimate an event's energy based on the total number of hit PMTs. This means that there is an uncertainty value associated with the reconstructed energy of each event, which was used as an extra criterion for selecting well reconstructed events.

The improvements in energy resolution for the FTK fitter over RSP are illustrated in figure 3.3. Figure 3.4 shows the resulting reduction in the number of ^{214}Bi and ^{208}Tl events leaking into the LETA analysis window.
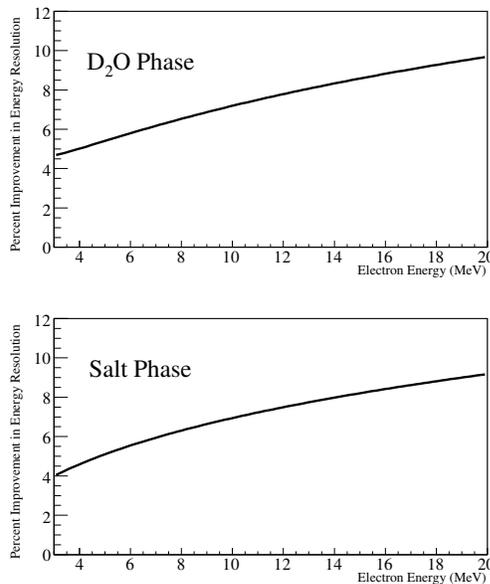
Figure 3.3: The percentage improvement in energy resolution between FTK and RSP in the D_2O and salt phases. Figure courtesy of M. Dunford.

Position Reconstruction

Two independent algorithms were utilised in the LETA analysis to reconstruct event positions within the detector. The path fitter, FTP, is a maximum likelihood reconstruction algorithm that incorporates time and angular information to fit for event position and
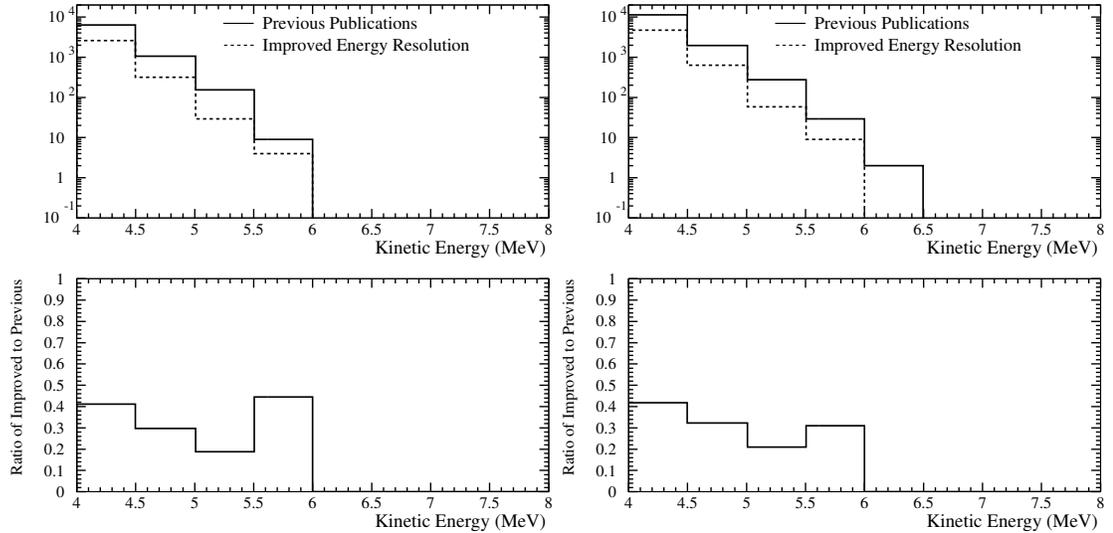
Figure 3.4: The reduction of ^{214}Bi (left) and ^{208}Tl (right) contamination in the analysis window resulting from the improved energy resolution. The top figures show the energy spectra using RSP (solid line) and FTK (dashed line) and the lower figures show the ratio of the two. Figures courtesy of M. Dunford.

direction [52]. FTP takes its seed from the quad fitter. The quad fitter takes each set of four PMT hits and uses the charge and time information to calculate a vertex. It then finds the region of highest density in this series of spacetime points. This results in poor resolution but very small tails - it is almost guaranteed to be in the correct vicinity. However, it can easily reconstruct outside the PSUP due to the broad resolution. FTP rejects any event with a seed vertex outside the PSUP and so has a relatively high failure rate.

The FTU fitter reconstructs vertices based solely on the time at which inward looking PMTs fired [53]. A full analysis of the accuracy of position reconstruction for both data and Monte Carlo events is presented in chapter 5, along with a discussion of the application of the resulting uncertainties to the final neutrino flux extraction.

Analysis Paths

Two independent analysis paths were defined for LETA, to reconstruct both the position and the energy of an event using independent methods, to act as checks on each other. The FTP position fitter was combined with the FTK energy fitter, and FTU was similarly combined with RSP. Henceforward, these two analysis paths will be referred to as the FTK/FTP path and the RSP/FTU path.

3.3.3 Selection Criteria

The LETA data set contains a large number of background events in the signal region, consisting of two main types. Physics backgrounds are events due to radioactive decays, as discussed in section 3.2. The decay of uranium and thorium daughters throughout the detector leads to a steep background wall at low energies, which is a particular issue for LETA. The second type of background present comes from events caused by phenomena related to the detection system itself, rather than from physical interactions within the detector. Such instrumental backgrounds include electronic pickup and static discharge in the PMTs.

In order to select neutrino events, a set of rejection criteria, referred to as data-cleaning cuts, were developed to remove as many of these background events as possible from the data set without substantial impact on the signal. These cuts are described below. Although these cuts were tuned very carefully, it is not typically possible to achieve perfect separation of signal and background events. A full analysis of the number of signal events lost to the cuts, known as the 'sacrifice', is presented in chapter 6. The measurement of the sacrifice of these cuts is important since the uncertainty on the sacrifice translates directly into an uncertainty on the final extracted neutrino fluxes.

Energy and Fiducial Volume Cuts

As described in section 3.2, the steep background wall limits the energy to which a neutrino spectrum can be studied. Although it would theoretically be possible to extract the spectrum at lower energies, the background event rate swamps the signal to such an extent that the uncertainties would dominate and no additional information would be gained. Therefore, a low energy threshold was set on the effective electron kinetic energy at $3.5\,\mathrm{MeV}$. Between 3.5 and $4\,\mathrm{MeV}$, background events dominate over the neutrino signal by a factor of 10; this low energy information helped to normalise the number of background events in the final extraction, reducing the uncertainty on the neutrino fluxes. For the purposes of the analysis presented in this thesis, the high end of the signal window was set at $11.5\,\mathrm{MeV}$. Again, it would be possible to extract information above this energy but very few neutrino events remain at higher energies so the statistical uncertainties dominate and very little information can be obtained.

To further define the signal window, cuts were placed on the reconstructed position of events. Although the D_2O volume extends to a radius of $600\,\mathrm{cm}$, background events from the rest of the detector leak into this volume. The radial distribution of these external backgrounds falls off exponentially, so by defining a fiducial volume inside which the analysis was performed, a large number of these events were rejected from the data set. In addition, the process of reconstructing event positions is less well understood near the

AV. Figure 3.5 shows the Monte Carlo predicted distribution of signal events and external backgrounds as a function of radial position within the detector. As for all LETA analyses, the radial parameter used in this figure is the volume-weighted, normalised parameter $\rho = (\frac{R}{R_0})^3$ where R is the radius of the event and R_0 is the radius of the AV. The fiducial volume was chosen to be at $550\,\mathrm{cm}$, $\rho \sim 0.77$, as shown on the figure. This volume defines the number of target deuterons for neutrino detection.
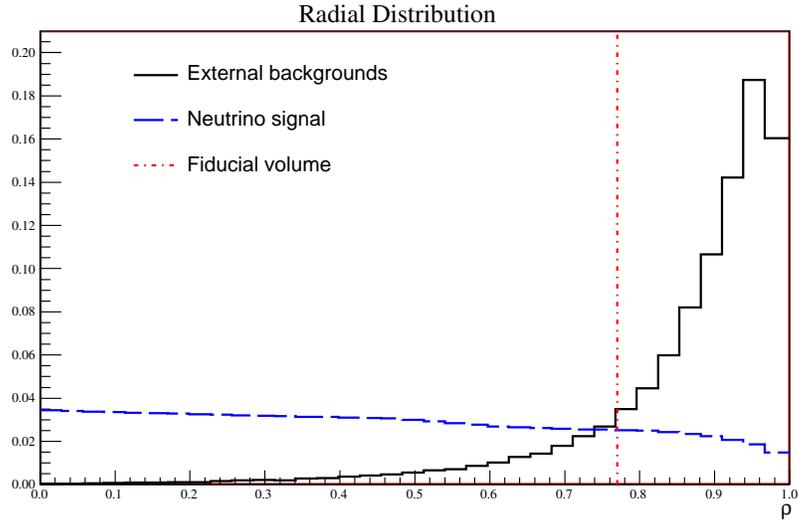
Figure 3.5: The normalised radial distribution of signal (CC+ES+NC) and external background events predicted by Monte Carlo simulations and the choice of fiducial volume.

Instrumental Cuts

Above energies of approximately $5\,\mathrm{MeV}$, the event rate was dominated by these events. Roughly ten neutrino events were expected per day in the detector, whereas the instrumental background rate was typically on the order of one per minute. Therefore, it was vital to understand these events and remove them from the data set for analysis.

Primary contributions to the instrumental background rate include static discharge inside the PMTs, generating events known as flashers which, although rare for any single PMT, become a significant contribution when integrated over the entire array. Neck events occurred when light was emitted from the neck region of the AV. Electronic pickup due to activity near the electronics crates occasionally caused false PMT hits in several channels at once.

These events have characteristics that are distinguishable from Čerenkov light and so can be identified and removed based on an analysis of the charge and timing distributions of the triggered PMTs, the spatial distribution of PMT hits and/or the firing of specific

veto tubes. For example, electronic pickup events are characterised by several channels with integrated charge values near the pedestal level. A number of cuts were developed to remove these instrumental backgrounds; these are fully described in [54]. Each one returns a simple binary decision, the result of which was stored in a Data Analysis Mask Number (DAMN) bank. A DAMN mask can be applied to any set of data in a bitwise manner to select only those events passing specific cuts.

High Level Cuts

A second set of cuts was utilised to further select signal events in the data set. A limited number of these *'high level cuts'* were used in previous publications; these were retuned and significantly extended for the purposes of the LETA analysis in order to deal with the increased number of radioactive background events [47]. In particular, this set of cuts targets external radioactive decay events originating in the light water or on the PSUP that misreconstruct inside the fiducial volume. Events are selected based on specific characteristics such as energy, isotropy and timing. It was found that different cuts were more effective on events reconstructed using different fitting algorithms and therefore two sets of cuts were developed - one for the FTK/FTP analysis and one for the RSP/FTU analysis.

A full description of all the high level cuts is given in [47]. An example of these cuts is that based on the isotropy of an event. Previous SNO analyses made use of two isotropy parameters: θ_{ij} and β_{14}. θ_{ij} is defined as the average angle between all hit PMTs in an event, relative to the fitted vertex. β_{14} was developed specifically for separating neutron and electron-type events in the salt phase [45]. β_{14} is a linear combination of the average value of the first and fourth Legendre polynomials of the cosine of the angle between each pair of hit PMTs in an event. This provides a tool for distinguishing external background events, since the light detected from events occurring outside the acrylic vessel is often less isotropic than that from events within the fiducial volume. Historically, θ_{ij} was used for D_2O phase analyses and β_{14} was used in the salt phase. For the LETA analysis, data sets from the two phases were combined, with the same cuts applied to each. It was found that, in this case, θ_{ij} was more effective for the FTK/FTP analysis whereas a cut on β_{14} was better for the RSP/FTU analysis path.

The result of the new set of high level cuts was to reduce the level of external background contamination by 25-30% over the old cuts, while increasing the loss of signal events by only ~2% [47].

3.4 Analysis Goals and Method

The goals of the LETA analysis are to extract the integral neutrino fluxes and the CC and ES spectra, across two phases of SNO data, to a lower energy threshold than that previously reported by any water Čerenkov experiment. In order to achieve a high precision result, many sources of systematic uncertainty were revisited and improvements were made to the modelling of the detector. The resulting reduction in systematic uncertainty, along with the increase in statistics, should result in the most precise measurement of the fluxes ever reported and, correspondingly, a more accurate evaluation of the MSW mixing parameters. The extraction of the CC spectrum to such a low energy should also allow for a test of predicted distortions in the incident ν_e energy spectrum.

These goals raised many new challenges, as discussed in the preceding sections. In particular, the lowering of the energy threshold significantly increased the number of background events reconstructing in the analysis window. In order to produce a high precision measurement, the three neutrino interaction types must be reliably distinguished both from each other and from these background events.

SNO detected events by observing the Čerenkov light produced by a recoil or Compton-scattered electron. Therefore, it is impossible to determine which type of interaction generated each individual event *i.e.* whether the electron came from a CC, ES, NC or even a background interaction. Hence, to extract values for the neutrino fluxes, a maximum likelihood fit was used to determine the most probable constituents of the full data set. Differing properties of the electrons produced in each type of interaction were exploited to separate the individual signal types. The maximum likelihood technique is described in detail in chapter 7; the main constituents of the fit are described below.

3.4.1 Signal Separation

For the purpose of this analysis, four observables were selected as the most powerful tools for event separation. These observables were: the reconstructed event energy (T_{eff}), the direction of the event relative to the Sun's location ($\cos\theta_\odot$), the position of the event in the detector (ρ) and the isotropy of the observed light (β_{14}, as discussed in section 3.3.3).

The energy spectra would be a strong tool in separating the different signal types. However, one of the aims of LETA was to increase SNO's sensitivity to the MSW-predicted turn-up in the neutrino survival probability, which would cause distortions in the energy spectrum. Both the CC and ES spectra are sensitive to this effect to some degree, so constraining the spectra in the fit would place assumptions on the underlying physics. Instead, both the CC and ES spectra were allowed to vary in the fit and only the NC spectrum was fixed, since the free neutron produced in this interaction has no memory of

the original neutrino energy.

The $\cos\theta_\odot$ distribution is a powerful tool for distinguishing ES events in particular, since the scattering of ν_e from the Sun results in electron events whose direction is strongly correlated with the location of the Sun. The CC also displays a weaker but noticeable correlation of $\sim 1 - \frac{1}{3}\cos\theta_\odot$. The γs generated by the capture of neutrons from NC interactions are produced in a random direction and, therefore, have no correlation with the Sun's location.

The radial position of events within the detector yields a weak separation between the three neutrino interaction types, but a much more powerful level of discrimination when external background events were introduced into the fit (see section 3.4.3). The CC and ES events occurred uniformly within the detector and, hence, their distributions are relatively flat in ρ. The neutrons produced in the NC interaction were generated uniformly, but those produced near the edge of the volume were more likely to escape into the H_2O region, where neutron capture produced γs of lower energy, which were therefore less likely to be detected. In addition, the cross section for neutron capture on the acrylic of the AV is very high and so the radial profile of NC events falls off at the edge of the volume. This effect is more noticeable in the D_2O phase data, since the neutron capture efficiency on deuterium is lower than on chlorine and, hence, the neutron mean-free path was longer in the D_2O phase than in the salt phase.

CC and ES events produce single electrons and, hence, are fairly anisotropic and yield a correspondingly high value for the isotropy parameter, β_{14}. The β_{14} distributions show small differences due to the different energy spectra of the two event types, which affects β_{14} through the known correlation between the energy and isotropy of an event (discussed in section 7.4.4). In the D_2O phase, neutrons captured on deuterium, resulting in a single γ of roughly $6.25\,\mathrm{MeV}$. This γ could Compton scatter an electron, which could, in turn, produce Čerenkov light if it was above threshold. The γ generally lost most of its energy to one electron and was less likely to scatter twice with high enough energies to produce two observable electron events, so the isotropy of NC events in the D_2O phase tends to be similar to that of the CC and ES events. In the salt phase, the neutrons captured predominantly on chlorine atoms, resulting in multiple, isotropically emitted γs. These could scatter multiple electrons and, hence, produce a more isotropic hit pattern, giving a lower value of β_{14}. Therefore, β_{14} provides a useful tool for separating neutron and electron-type events, particularly in the salt phase.

The Monte Carlo simulation was used to generate the expected distributions in each of the four observables previously described, for each interaction type in each phase, assuming an incident flux of pure ν_e, with no oscillation or spectral distortion. These distributions are shown in figures 3.6 and 3.7. Each distribution has been normalised to

1.0, except the $\cos\theta_\odot$ distributions for CC and NC events which have been scaled by a factor of 10 in order to show them clearly against the scale of the ES distribution.
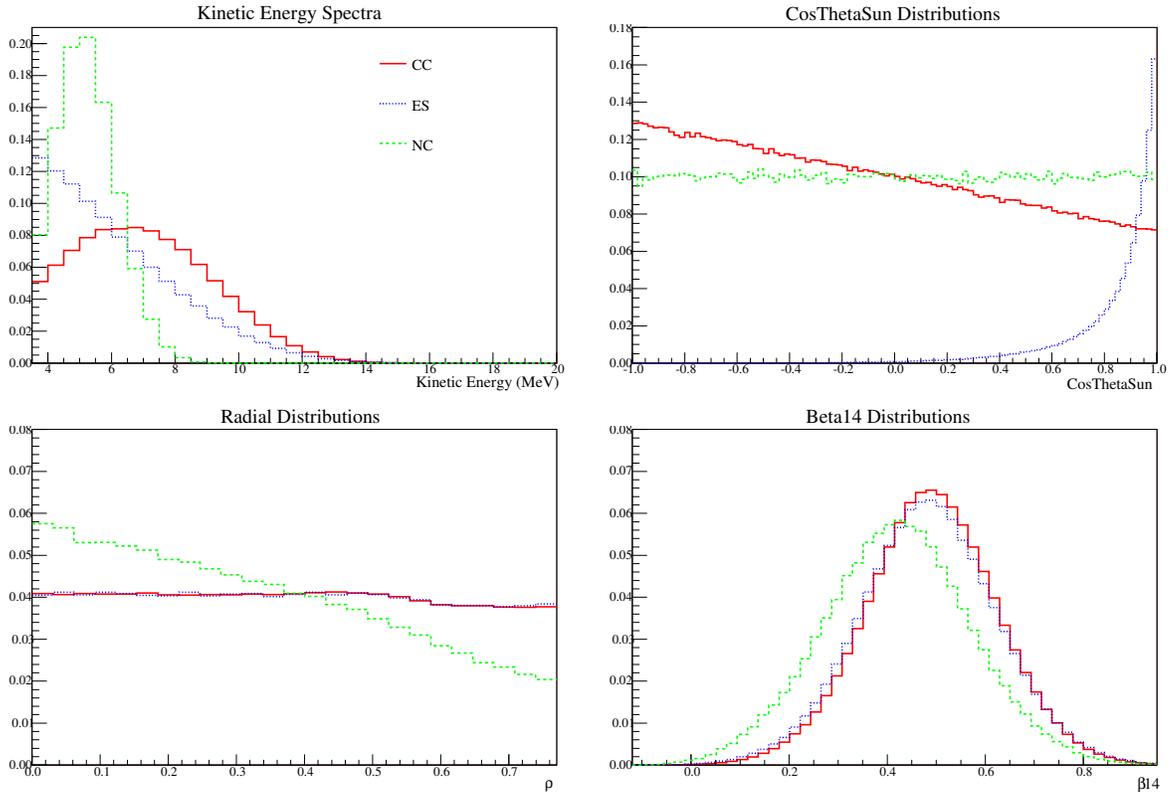
Figure 3.6: Normalised distributions of the observables used to separate event types in signal extraction for the three neutrino interaction types in the D$_2$O phase. The $\cos\theta_\odot$ distributions for CC and NC events have been scaled by a factor of 10 for clarity.

3.4.2 Multiphase Fitting

A new aspect of the LETA analysis was the simultaneous fitting of two independent data sets from two phases of the SNO experiment. The flux of neutrinos from the Sun is assumed to be constant over time and so the number of neutrino events in the two phases cannot vary independently. Full details of the constraints applied between the phases are given in chapter 7. Constraining the two data sets to contain the same rate of incident solar neutrinos makes full use of the all information available, taking into account all correlations and dependencies at once. For example, the different information available in the two phases, such as the differing isotropy distributions, can be used to tighten the constraint in a joint fit more than that of any single phase.
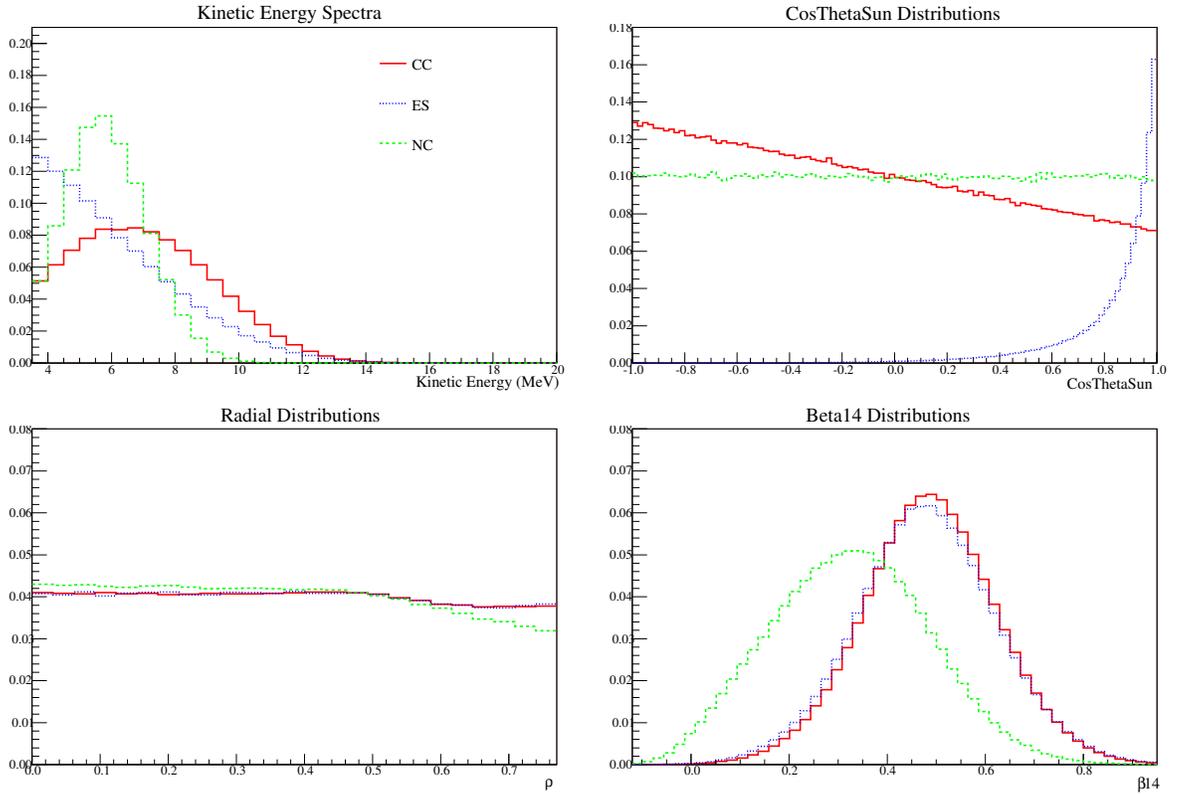
Figure 3.7: Normalised distributions of the observables used to separate event types in signal extraction for the three neutrino interaction types in the salt phase. The $\cos\theta_\odot$ distributions for CC and NC events have been scaled by a factor of 10 for clarity.

3.4.3   Incorporating Background Events

In previous SNO publications, the energy threshold was high enough that contamination from low energy background events was very low and, therefore, these events were not included in the fit. As discussed in section 3.2, the lowering of the analysis energy threshold caused a significant rise in the number of background events observed in the analysis window. Therefore, it was necessary to include all the possible sources of low energy background events in the fit. The necessary modifications to the maximum likelihood fit to incorporate these additional event types are described in chapter 7.

The sources of possible background events were discussed in section 3.2. Trace quantities of naturally occurring radioisotopes were found in all components of the detector. The dominant decays observed in the LETA analysis window are those from the ^{238}U and ^{232}Th decay chains, in particular the decays of ^{214}Bi and ^{208}Tl. These produce both Čerenkov events from β-γ decay branches and neutron events due to photodisintegration of the deuterium nucleus by γs above 2.23 MeV. In the salt phase, neutron capture on

^{23}Na resulted in a low level of ^{24}Na in the detector. ^{24}Na decays to ^{24}Mg, producing a low energy β and a 2.75 MeV γ, which could photodisintegrate a deuteron. Therefore, this contributed some level of both electron and neutron-type background events. In addition to these backgrounds, a small number of neutron events were observed on the surface of the AV from (α,n) reactions on isotopes of C and O that accumulated on the surface of the acrylic during construction.

The backgrounds that were included in the full likelihood fit were therefore: ^{214}Bi and ^{208}Tl decays in the D$_2$O volume, in the bulk of the AV itself and in the H$_2$O volume for each phase of data taking; PMT β-γ events (^{214}Bi and ^{208}Tl decays in the PMT glass) in each phase; neutron events on the surface of the AV occurring in each phase and ^{24}Na backgrounds in the salt phase only. This gives a total of 17 additional signal types that were included in the signal extraction procedure. These backgrounds are summarised by phase and detector region in table 3.1.

Detector Region	D$_2$O Phase	Salt Phase
D$_2$O volume	Internal ^{214}Bi Internal ^{208}Tl	Internal ^{214}Bi Internal ^{208}Tl ^{24}Na
Acrylic vessel	Bulk ^{214}Bi Bulk ^{208}Tl Surface neutrons	Bulk ^{214}Bi Bulk ^{208}Tl Surface neutrons
H$_2$O volume	External ^{214}Bi External ^{208}Tl PMT β-γs	External ^{214}Bi External ^{208}Tl PMT β-γs

Table 3.1: The sources of physics-related background events in the LETA analysis.

The Monte Carlo was used to generate the distributions in each of the four observables for these background events. The distributions are shown in figures 3.8-3.11. The power of the radial parameter in particular for separating out the external backgrounds is very clear, since these events have a very distinct, steep distribution in ρ. The internal backgrounds have very similar distributions to the CC signal and, therefore, are more difficult to distinguish. This is discussed further in chapter 7.

Constraints on Background Levels

During data taking, several radioassays were performed to measure the concentrations of ^{238}U and ^{232}Th in the D$_2$O and H$_2$O regions, as discussed in chapter 2. The results were used to place constraints on the expected number of background events in the analysis window. During the salt phase, there was a leak in the assay system used to measure the
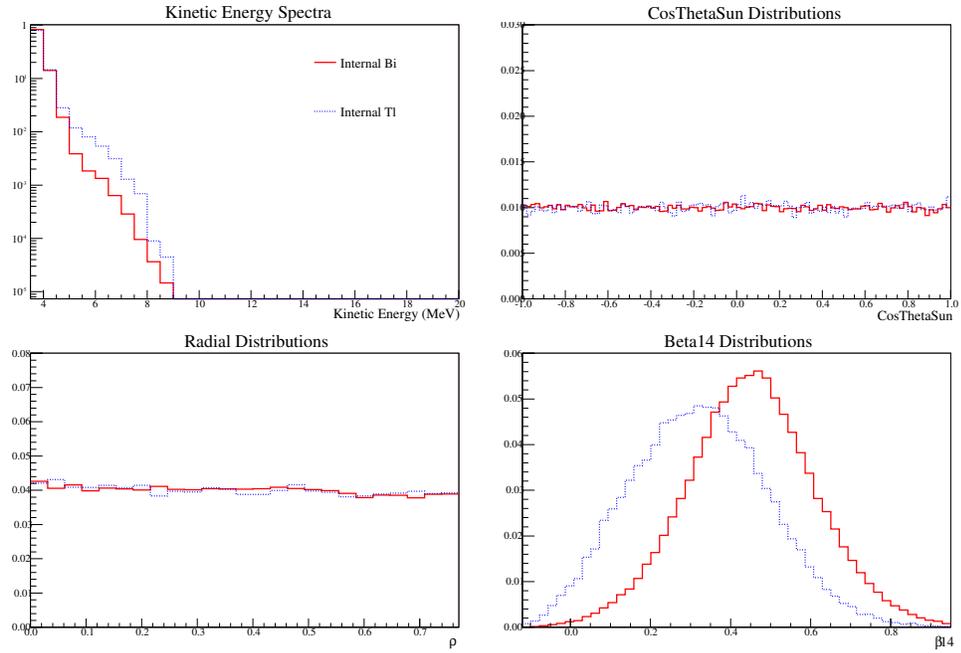
Figure 3.8: Distributions of the observables used to separate event types in signal extraction for internal background events in the D$_2$O phase, normalised to 1.0.
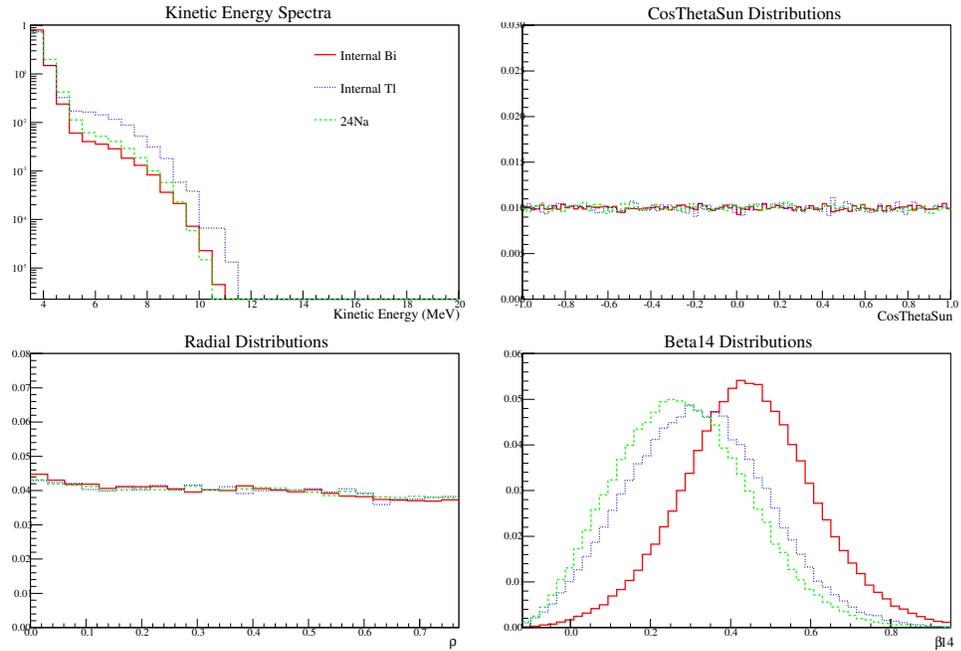
Figure 3.9: Distributions of the observables used to separate event types in signal extraction for internal background events in the salt phase, normalised to 1.0.
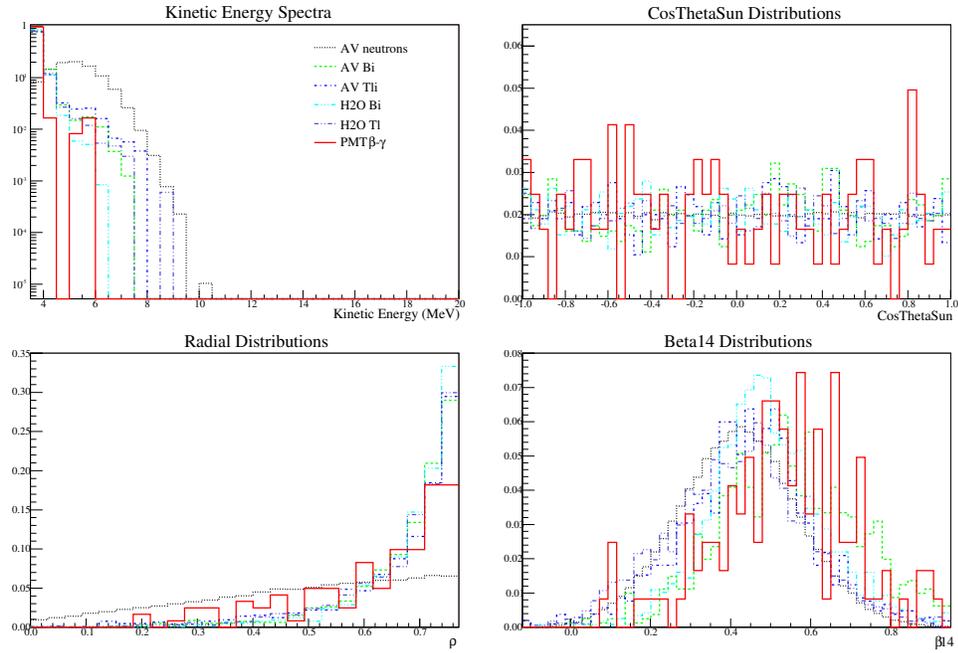
Figure 3.10: Distributions of the observables used to separate event types in signal extraction for external background events in the D_2O phase, normalised to 1.0.
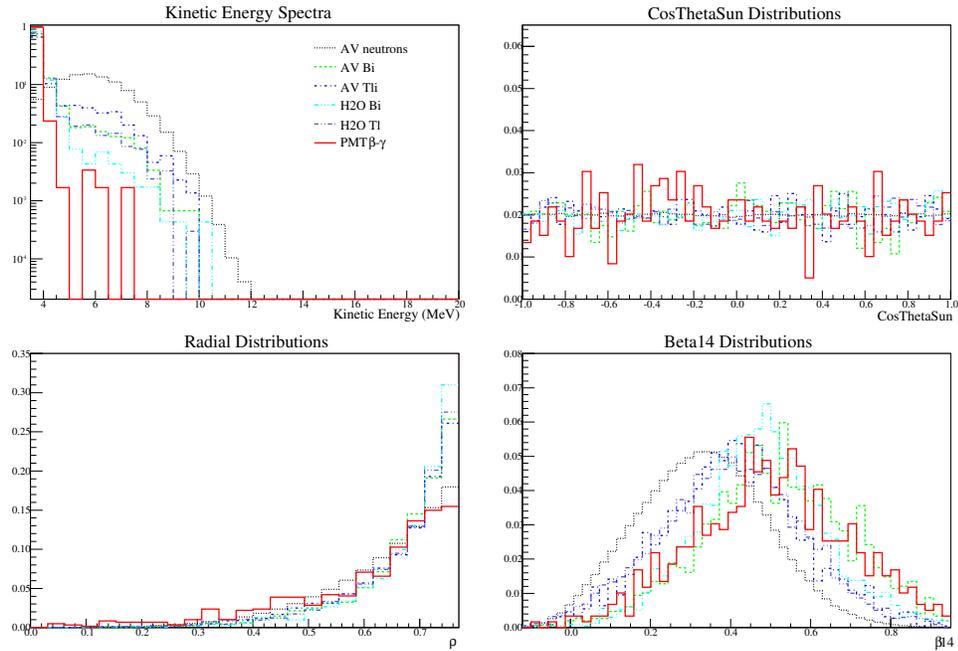
Figure 3.11: Distributions of the observables used to separate event types in signal extraction for external background events in the salt phase, normalised to 1.0.

^{238}U chain contamination. This was not discovered until after data taking had ended and so there is no accurate constraint on the level of ^{238}U in the D$_2$O volume in the salt phase. Loose limits exist based on secondary assay techniques, but these were found to be too wide to have any impact on the signal extraction procedure and so were not implemented. The results of the assays are given in table 3.2.

Phase	Isotope	D$_2$O Volume (g/g of D$_2$O)	H$_2$O Volume (g/g of H$_2$O)
D$_2$O	^{238}U	$1.01^{+0.34}_{-0.20} \times 10^{-14}$	$29.5 \pm 5.1 \times 10^{-14}$
	^{232}Th	$2.09 \pm 0.21(\text{stat.})^{+0.96}_{-0.91}(\text{syst.}) \times 10^{-15}$	$8.1^{+2.7}_{-2.3} \times 10^{-14}$
Salt	^{238}U	—	$20.6 \pm 5.0 \times 10^{-14}$
	^{232}Th	$1.76 \pm 0.44(\text{stat.})^{+0.70}_{-0.94}(\text{syst.}) \times 10^{-15}$	$5.2 \pm 1.6 \times 10^{-14}$

Table 3.2: ^{238}U and ^{232}Th concentrations in the D$_2$O and H$_2$O volumes, determined from *ex-situ* radioassays in the D$_2$O and salt phases.

These results are given in terms of the grams of contamination of each isotope per gram of detector medium. The mass of D$_2$O and H$_2$O is known, so this can be straightforwardly converted into the total expected mass of each isotope present in each detector region. The decay rate per gram of isotope is given by:

$$\frac{N_A \times ln(2)}{A \times T_{\frac{1}{2}}} \tag{3.1}$$

where N_A is Avogadro's number, A is the isotope's mass number and $T_{\frac{1}{2}}$ is its half-life. The half-lives of ^{238}U and ^{232}Th are 1.41×10^{17} s and ^{232}Th is 4.45×10^{17} s, respectively. Multiplying the decay rate per gram of the isotope by the total mass gives the total expected decay rate of ^{238}U and ^{232}Th in each region of the detector. Under the assumption of secular equilibrium, the rate of daughter decay events is equal to parent decays so, allowing for branching fractions, this gives the expected rate of ^{214}Bi and ^{208}Tl decays in the detector. 99.99% of ^{238}U goes to ^{214}Bi, so the rate of ^{214}Bi decays is equal to the calculated rate of ^{238}U. However, the branching fraction of ^{232}Th to ^{208}Tl is 0.36, meaning that only 36% of ^{232}Th decays result in a ^{208}Tl daughter, so the ^{232}Th rate must be multiplied by 0.36 to give the expected ^{208}Tl rate.

These ^{214}Bi and ^{208}Tl rates were incorporated into the final signal extraction as constraints on the number of background events in the D$_2$O and H$_2$O regions. The exact nature of the constraint is detailed in chapter 7. In addition to these *ex-situ* constraints, *in-situ* analyses performed by A. Hallin and C. Sims were used to constrain the expected levels of ^{24}Na background in the detector [35],[55]. This constraint was calculated to be 392 ± 117.6 events across the livetime of the salt phase.

3.4.4 Systematic Uncertainties

The Monte Carlo was used to generate the PDFs for each signal type. Any uncertainty associated with the shapes of those distributions must be measured and propagated through the signal extraction procedure as a systematic uncertainty.

In the past, the effects of systematic uncertainties on the extracted neutrino fluxes were evaluated by applying distortions to the distributions in each observable for each signal type and re-extracting the neutrino fluxes to determine the effect. This method did not allow for any correlations between different systematic effects, since each one was applied individually. It also did not allow the data itself to add constraints on the size of the systematic shift. Although the derivation of each systematic uncertainty was revisited for LETA, an additional aim of the analysis was to formulate the final neutrino flux extraction in such a way that further constraints on these values could be obtained from the data. This is discussed in more detail in chapter 7.

3.5 Summary

The LETA analysis aims to extract the CC spectrum down to an effective electron kinetic energy of $3.5\,\mathrm{MeV}$ and to improve upon the precision of SNO's previous measurements of the solar neutrino fluxes, resulting in the best possible constraints on the parameter space. Many aspects of the detector modelling have been improved and many new analysis tools have been developed in order to meet this goal.

The work presented in this thesis focuses on improving the modelling of the detector and our understanding of the associated uncertainties in order to perform a full extraction of the neutrino fluxes down to this lowered energy threshold.

Chapter 4 describes improvements made to the modelling of the angular response of the PMT and light concentrator arrangement in the simulation code. This work resulted in a significant reduction in the radial bias previously observed in the energy estimators, rendering the inclusion of that part of the energy scale uncertainty unnecessary. Chapter 5 presents a full analysis of the accuracy of the position reconstruction algorithms. The uncertainties associated with position reconstruction are measured and their application to the neutrino flux extraction discussed, with particular focus on any energy dependence of the uncertainty and its impact on a spectral measurement. Chapter 6 contains a full study of the acceptance of the data-cleaning cuts employed for the LETA analysis. The application of these cuts is expected to result in some level of sacrifice, where signal events are removed from the data set along with the background events. Since the Monte Carlo distributions were used to define what is expected in the LETA analysis window, any factors affecting the number of events observed in that window must be fully understood.

Uncertainties on the sacrifice translate directly into uncertainties on the extracted neutrino fluxes. In particular, the energy dependence of these uncertainties has a direct impact on the extraction of a neutrino spectrum.

Chapter 7 describes the procedure for extracting the neutrino fluxes from the full LETA data set. The use of multiple phases and the incorporation of background events into the fit is discussed and a verification of the PDF shapes used in the fit is presented, along with extensive testing of the maximum likelihood fitting method and a discussion of the handling of systematic uncertainties. Chapter 8 presents the final flux and spectrum results for the FTK/FTP analysis path. Also presented is an interpretation of these results and their impact on the parameter space. A rigorous and powerful calculation of the contour plot that makes direct use of the likelihood calculation involved in the signal extraction gives the most accurate constraints on the values of the MSW parameters.

Chapter 4

Angular Response

A full understanding of the optical response of the SNO detector is a crucial part of estimating the energy of events occurring within the detector. The amount of Čerenkov light generated by an event is directly related to the deposited energy. The number of photons that triggered PMTs within a specific time window is used to reconstruct that energy, so it is important to understand any factors that could affect the fraction of generated photons that were detected. One of these factors is the optical response of the PMTs and light concentrators themselves. The detection efficiency of the PMT and light concentrator arrangement has a strong dependence on the angle at which a photon enters the concentrator region. Degradation of the concentrators can cause the shape of this angular response to change with time. This chapter describes bench-top measurements of the angular response of the PMT and light concentrator setup. Also presented is a full study and optimisation of the modelling of the response in the Monte Carlo code, including incorporating degradation due to aging into the simulation.

4.1 PMT and Light Concentrator Set-up

Energy estimation within the SNO detector involves using the charge deposited in the PMTs to determine the number of Čerenkov photons generated by an individual event. A number of factors come into play in this calculation, including the optical response of the detector. Fully understanding this response and accurately modelling it in simulation code is therefore a vital ingredient in the process of reconstructing event energies.

Each PMT-concentrator arrangement (hereafter referred to as the 'PMT bucket' or just the 'bucket') consisted of an eight inch Hamamatsu R1408 PMT and a light concentrator, designed to sit flush against the PMT glass. The concentrators were designed to focus light on to the active area of the PMT's photocathode, thus increasing the light collection efficiency and hence improving both the position and energy resolution of the

detector. A bonus of using concentrators was that they also shielded the PMTs from light due to background events associated with either neighbouring PMTs or the PSUP. Figure 4.1 shows the geometry of the PMT bucket. The concentrator shape and the PMT bucket geometry is discussed in more detail in [56] and [57]. The reflective surface was achieved using 18 dielectric coated aluminium segments, known as 'petals', mounted on a moulded plastic support. A full description of the design and modelling of the concentrators is given in [58].
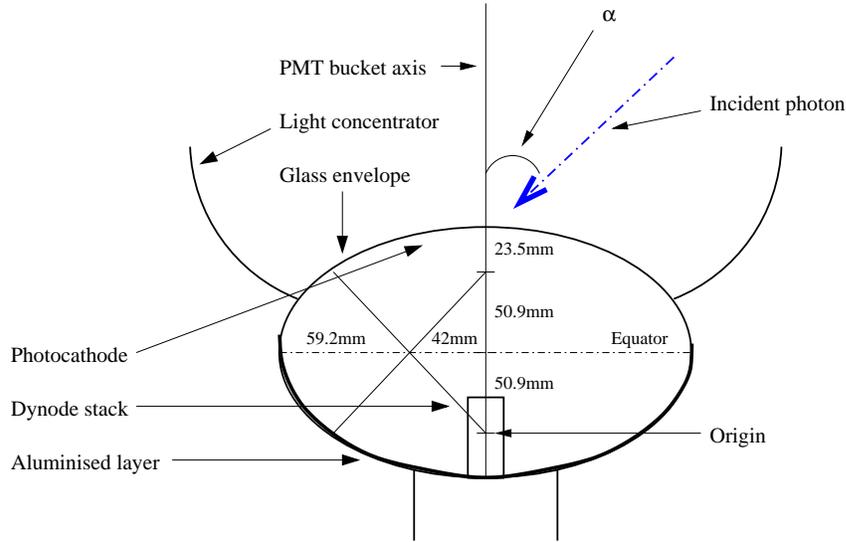
Figure 4.1: A schematic diagram of the PMT and light concentrator setup.

The response of the PMT and concentrator to incident light is measured in terms of the angle at which a photon enters the PMT bucket, shown in figure 4.1 as α. This angular response can be affected by several factors: the detection efficiency of the PMT varies with the position at which a photon strikes it due to variations in the photocathode response across the front face of the PMT; the reflectivity of the concentrator petals varies along their length and the reflective coating on the petals degraded due to immersion in water, which could cause the reflectivity to vary in a time dependent fashion.

It is important for the simulation to accurately model whether a photon striking the front face of the PMT will cause a count to be registered by the electronics. Lay [57] and Lyon [59] first investigated the angular response in both air and water in order to understand the PMT bucket behaviour and to model it in the Monte Carlo code. The original response in the code was based on these measurements. During the transition between the D_2O and salt phases of data taking in SNO, calibration data taken within the detector showed a marked decrease in the light collection efficiency. This was thought to be caused by increased corrosion of the concentrator petals due to a noticed fluctuation in the

temperature of the water surrounding the PMTs. As a result, further measurements were taken, which highlighted a discrepancy in the recorded response from that predicted by Monte Carlo simulations [60], [61]. Detailed studies were performed, which demonstrated that the shape of the response was changing in the detector, not just the overall collection efficiency, in a way that was not reproduced in the simulation [62].

The first section of this chapter involves a verification of previous bench-top measurements of the angular response and an investigation into the effects of concentrator aging. In subsequent sections, the modelling of the angular response in the Monte Carlo code is optimised to better reproduce the calibration data taken in the detector. In addition, a degradation is introduced in the shape of the response to model the change of the concentrator reflectivity over time.

4.2 Experimental Measurements of Angular Response

4.2.1 Apparatus

The apparatus used for these experiments is shown in figure 4.2. The Čerenkov light was generated from a Strontium-90 source housed in a metal tube, mounted on the underside of a window in the base of the apparatus. The source, developed by Boardman [63], consisted of a drop of ^{90}Sr liquid sealed in an acrylic sphere. ^{90}Sr is a β-emitter, generating electrons energetic enough to produce Čerenkov light in the acrylic. An opaque, black mask attached to the window above the source prevented any upward going light reaching the PMT directly. The ^{90}Sr source was located at the focal point of a parabolic mirror, so that downward going photons would be reflected back up towards the PMT parallel to the vertical axis. The distance of the source and mirror from the PMT setup was chosen such that the total photon path length would ensure a rate of one photon at the PMT per registered event. This allowed a measurement of the single photoelectron peak. The PMT and concentrator arrangement was mounted on a pivot to allow it to be rotated with respect to the base of the apparatus, thus allowing a range of angles of incident light to be measured.

A mu-metal shield encased the PMT and concentrator setup. This shield was constructed from a material with a high magnetic permeability in order to protect the apparatus from any effect of the earth's magnetic field.

The PMT was supplied with a 2 kV DC voltage. The output was amplified by a factor of 10 and then passed through a discriminator to remove any background noise. The threshold setting of the discriminator was chosen to be 30 mV to cut out electronics noise but not interfere with the single electron peak [61]. The output of the discriminator was then connected to a dual scalar to count the number of hits registered on the PMT.
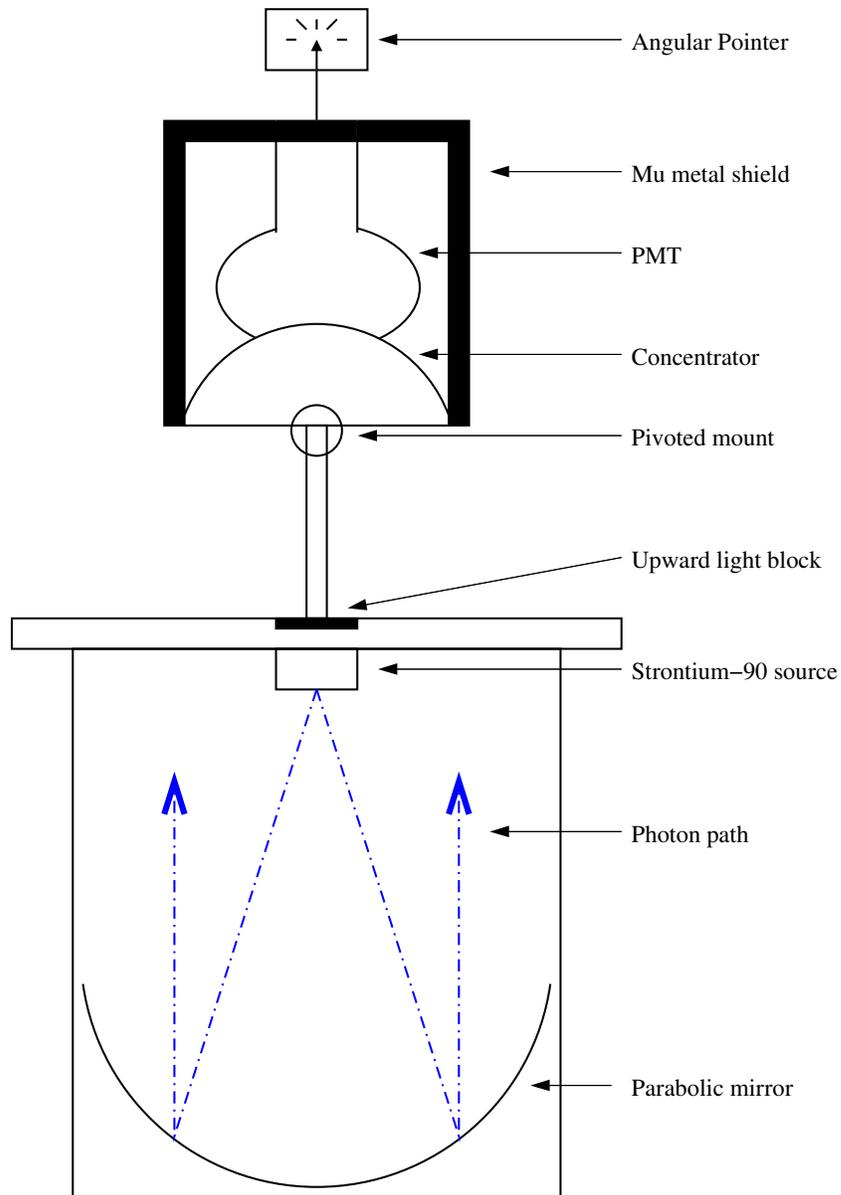
Figure 4.2: A schematic diagram of the experimental apparatus used to measure PMT and light concentrator angular response.

To check the output at each stage, an oscilloscope was connected to view the electron pulse from the PMT and the logic pulse from the discriminator.

4.2.2 Method

The PMT is very sensitive to light, so many precautions had to be taken to ensure that external sources of light did not affect the measurement of the signal from the source. A

dark room was set up, with the walls and ceiling painted entirely in a matt black paint. Any cracks in the ceiling through which light could intrude were sealed with black tape. A blind was fixed to the interior of the doorway and a heavy black curtain was hung across the outside to minimise possible sources of light leaks. Measurements were taken from an outer room. To ensure minimal light leakage, the main lights in that room were kept off and only a single red light bulb was used since the sensitivity of the PMT to wavelengths above 600 nm is significantly reduced (see section 4.3.1).

Any exposure to high light levels while the PMT is active can damage the electronics, so the voltage supply had to be switched off for each change that was made to the apparatus, such as rotating the PMT. The voltage was then ramped back up slowly and a few moments allowed for the PMT to settle and for the count rate to stabilise before further readings were taken. Each photoelectron generated in the PMT causes a cascade in the dynode stack; variations in the number of electrons produced at each dynode lead to a distribution of pulse heights. Care was taken to ramp the voltage up to the same value each time since the exact voltage supplied affects the multiplication at the dynodes and, hence, the gain of the PMT.

Even after blocking all possible sources of light leaks, a non-zero background reading was expected since thermal excitation at the single electron level causes some activation of the PMT. To allow a measurement to be made of the background count rate, a light blocker was fitted just beneath the base window of the apparatus to block all light from the source. The background rate could then be used to verify that the mu-metal shield was successfully cancelling any effect of the earth's magnetic field. A non-zero magnetic field would distort the path of the electrons in the dynode stack and, hence, affect the gain of the PMT. This effect would be dependent on the orientation of the dynodes relative to the field and, hence, on the orientation of the PMT. The background due to thermal excitation should be uniform and, so, any non-uniformity observed would indicate either a residual magnetic field or an unblocked light leak.

Measurements were taken at ten degree intervals to define the shape of the response curve. Each measurement was taken by recording the number of PMT hits above the discriminator threshold in a time window of 100 seconds. Three background counts were taken at each angle, then three signal counts and three further background counts as a stability check. The results at each angle were averaged and the background reading subtracted from the signal to give the final measurement. Since the shape was under investigation, not the overall light collection efficiency, the response was normalised to 1.0 at normal incidence. [*]

---

[*] Normal incidence is defined as light entering the PMT bucket at an angle of 0° to the vertical PMT axis.

Two light concentrators were employed for the investigations into concentrator aging. One was a new concentrator, still in its original condition. The second was aged in a test tank on site by being immersed in water at a temperature of 20°C for three years. This is equivalent to approximately six years in the conditions of the SNO detector (~10°C).

4.2.3 Results

To check the efficiency of the magnetic field shielding, two sets of background readings were taken on the apparatus in different orientations. First, the apparatus was aligned so that the dynodes were perpendicular to the direction of the earth's magnetic field, resulting in parallel magnetic and electric fields and, hence, no resultant force on the electrons. A full set of readings were taken and then the apparatus was rotated by 90° so that the dynodes were parallel to the earth's field, giving perpendicular E and B fields. If the magnetic shielding was not fully effective, this would result in a force on the electrons that would manifest itself as an asymmetry in the background count rates as the PMT was rotated about the polar axis.

Figure 4.3 shows the background readings taken at each angle for these two different orientations. The error bars represent the spread observed in individual readings. As can be seen, the two different orientations were entirely consistent, with no clear asymmetry apparent in one that was not seen in the other. This was taken as evidence that the magnetic shielding was effectively blocking the earth's field.
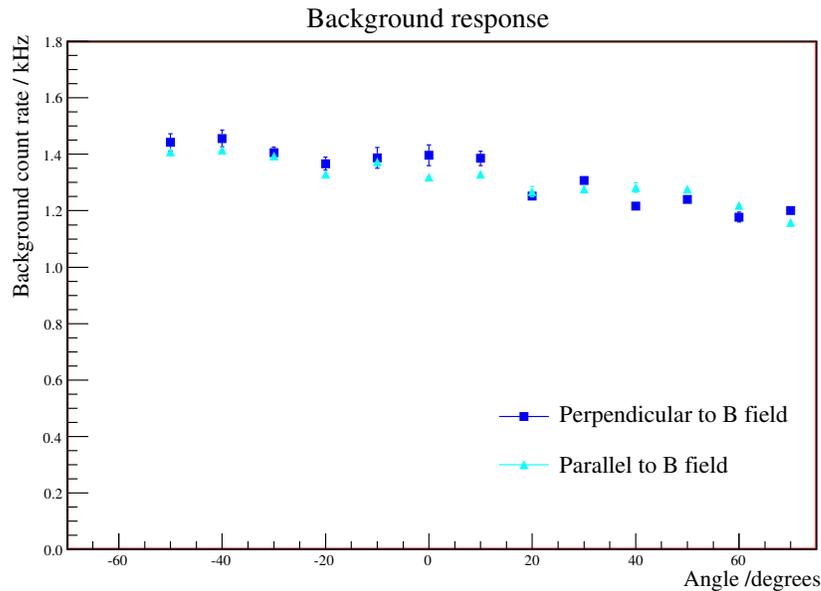
Figure 4.3: Background measurements taken using a SNO PMT and light concentrator.

In comparison to previous studies [61], the average background count observed was noticeably lower (1.4kHz in comparison to previous observations of 2kHz), suggesting that the construction of the dark room was more effective than previous methods of sealing the apparatus off from the light. However, the weak asymmetry that was observed in both sets of measurements still implied the presence of a residual light leak that had not been fully blocked, which affected the rates more as the PMT was rotated and its face exposed to a greater extent. The method of subtracting the background measurement from the signal count at each angle, as described above, was deemed sufficient to correct for this weak effect.

Once the stability of the background rates had been determined, the light blocker was removed so that the signal from the source could be measured. The first stage was to take a full set of readings using the same PMT as used in the previous tests to check the consistency of the results. The response was shown to agree with the previous measurements and to show the same disagreement from the Monte Carlo prediction.

A proposed explanation for this was a variation in the response across different PMTs. To test this, the measurements were repeated with a second PMT. The results are shown in figure 4.4. The discrepancy of the original PMT measurements from the Monte Carlo prediction is clear, whereas the second PMT response is in perfect agreement with the simulations. For the purpose of this work, the simulation was run with air surrounding the PMTs, instead of water as in the true detector. This allows a fair comparison of the experimental results to the Monte Carlo predictions.

Using the second PMT, investigations were then performed into the effect of concentrator aging. The response was measured for both a new and an aged concentrator and the results are shown in figure 4.5. The left insert shows a comparison of the response for the two concentrators across the full range measured (0-70°). The response of the aged concentrator is reduced at wider angles; in particular, the response is lower at the peak of the curve, around 40°. The right hand insert shows the response for both concentrators, comparing the positive and negative sides of the response curve. An asymmetry can be seen in the response of the aged concentrator that is not present in the new concentrator's response.

4.2.4   Conclusions

The work presented here has verified that the variation in response from one PMT to another is sufficient to account for the discrepancies seen between previous measurements of the PMT angular response. In addition, aging of the light concentrators results in a reduced response, particularly at larger angles, where the response is at its peak. This could account for the reduction in light collection efficiency observed between the two
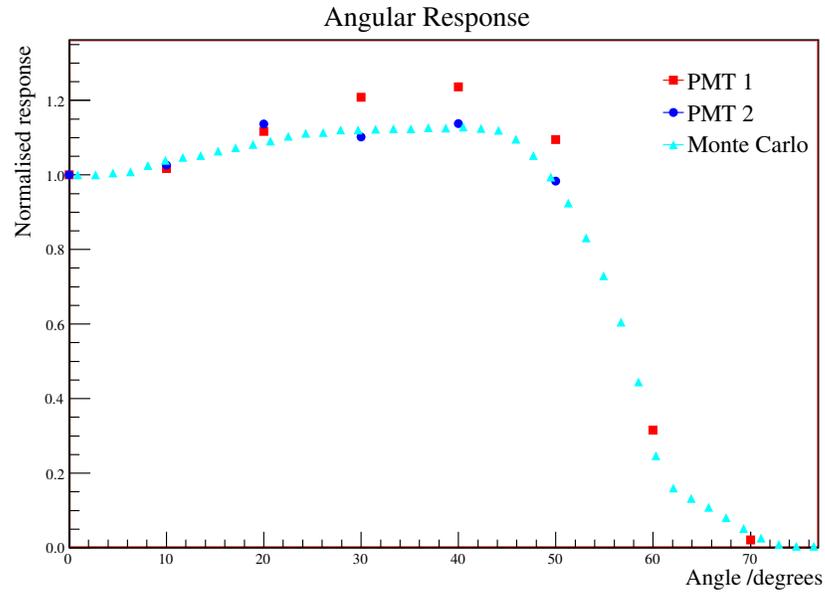
Figure 4.4: Angular response measurements taken using two SNO PMTs (with light concentrator) and a comparison to the Monte Carlo prediction.
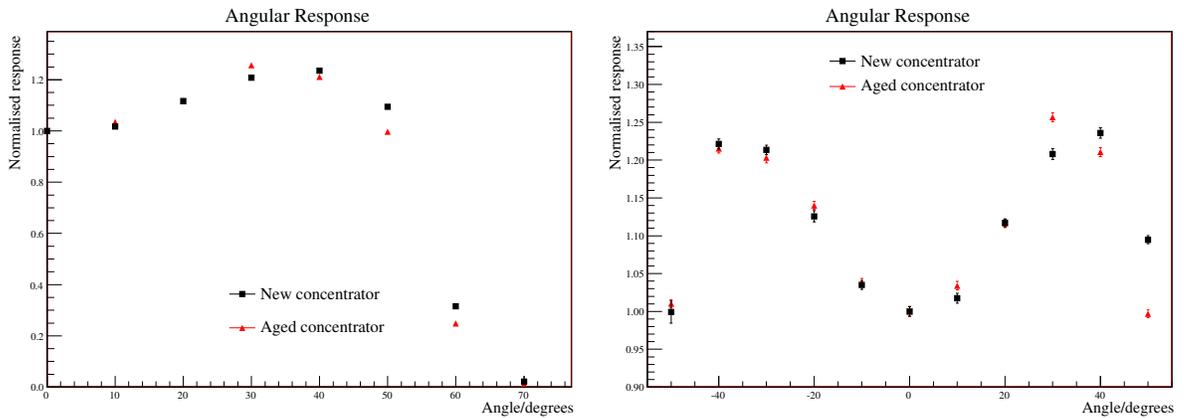
Figure 4.5: The effect of aging on the angular response of SNO light concentrators.

phases of data taking in the SNO detector under consideration. Concentrator aging has also been observed to be a non-uniform process, which could further impact the detector response.

4.3 Modelling the Angular Response in Simulations

The original modelling of the angular response in the simulation code was based on measurements made by Lay [57]. Calibration data taken in the detector showed that some modification of that response was required to accurately model the *in-situ* response of the PMTs. Accordingly, amendments were made that included adjusting the dynode stack reflectivity and incorporating a variable thickness for the photocathode across the face of the PMT [64]. Both these adjustments were based on bench-top measurements of the relevant physical properties. The response was still not accurately reproduced and there were no tunable parameters to correct it, so a modified collection efficiency (MCE) function was created that altered the response of the tube as a function of the vertical position at which a photon struck the photocathode [64]. In effect, this altered the angular response of the tube. This function was a purely empirical addition, designed to make the simulated response reproduce the calibration data as closely as possible.

The results were an improvement but still not a perfect match to the data. In addition, calibration data taken at different periods during data-taking showed a marked change in the response over time. The purpose of the work presented here was to optimise the response in the code to more closely reproduce calibration data and then to incorporate a model for the effect of concentrator aging.

For all the work presented here, the simulated PMT angular response was extracted by generating photons at fixed energies (according to the wavelength required) in the light water region of the detector, uniformly positioned on a spherical shell at a radius of 800 cm. The track information stored in the code was used to determine the angle at which the photon entered the PMT bucket and whether or not that photon resulted in a successful PMT hit. The response was defined as the fraction of photons entering the bucket that resulted in successful PMT hits, normalised to 1.0 at normal incidence.

The calibration data used in this work was obtained from laserball scans taken within the SNO detector [65]. Data was taken at six wavelengths (337, 365, 386, 420, 500 and 620 nm) at one time during the D_2O phase and at several different periods in the salt phase.

The optics data from the laserball only covers the region of the response up to 40°. The concentrators cause a cut-off in the PMT response at around 60-65°, so a different set of information was needed for the wide angles. At first, no constraint was put on the shape of the response above 40°, but this resulted in a very poor match in the response in the light water region of the detector, where these wide angles are more significant. Therefore, it was determined that bench-top measurements taken by Lay [57] would be used to model the shape of the response beyond 40°. Since these were *ex-situ* measurements, the overall light collection efficiency was not as directly applicable as the *in-situ* laserball calibrations.

The normalisation was therefore fixed by matching the response to the laserball response at the boundary. Large uncertainties were applied to the *ex-situ* region of the resulting response, in order not to overweight this region when tuning the simulation to match the shape of this response.

4.3.1 Tuning the Angular Response

Initially, the possibility of modifying either the petal reflectivity or the photocathode thickness was investigated. The existing model for the photocathode thickness was based on bench-top measurements made for a number of PMTs [64], so a physical reason would be needed to justify any alterations. The simulated petal reflectivity was also based on bench top measurements, but aging of the petals due to immersion in water had already been shown to alter their reflectivity. Therefore, the code was extended to modify the petal reflectivity according to various models. Investigations were performed into both a uniform degradation of the reflectivity across the surface and a linear degradation along the length of the petals. Preliminary test results showed that although an improvement was possible with this method, there was not enough scope to change the shape of the response sufficiently to match the data without impacting on other features of the detector response. Altering the petal reflectivity in order to reproduce the shape of the response in data significantly reduced the overall light collection; in particular, the late light response was affected more than direct light. In addition, the assumption that the concentrator petals were as new at the start of the D_2O phase of data taking in SNO meant it should not be necessary to degrade the reflectivity to reproduce the detector's response at this stage.

For these reasons, the MCE was targeted for any changes. This function was introduced purely as an empirical function to improve the simulation response. The form of the function is given in equation 4.1, where *apmt* is one of the dimensions of the inner side of the glass envelope and *zpm* is the height at which the photon strikes the PMT:

$$
\begin{aligned}
x &= (zpm - apmt)^4 \\
\text{MCE} &= (a_1 - a_2 e^{(b_1 x)}) e^{(b_2 x)} + a_3
\end{aligned}
\tag{4.1}
$$

This function modifies the collection efficiency of the PMT across the photocathode. Although this function depends on the height at which a photon strikes the PMT, previous work demonstrated that this was strongly correlated with the angle at which the photon entered the PMT bucket [64], so in effect this function modifies the angular response.

a_1, a_2, a_3, b_1 and b_2 are parameters of the function, which can be modified to improve the shape of the angular response. Initially, an iterative procedure was used to modify

these parameters. However, it was found that there was no obvious correlation between the changes introduced to these parameters and the effect on the shape of the response and, so, a more efficient method was required to find the best values. SNOMAN was modified to output data at the stage at which photons enter the optical response portion of the code, before they strike the photocathode. An existing piece of code [50] was then modified to read in the output from SNOMAN and calculate the angular response for a specific set of values of the five parameters in the MCE. The MINUIT minimisation package [66] was used to optimise the values of these parameters in order to minimise the χ^2 of the fit of the simulated response to calibration data. Once a best fit set of parameters was found, the full SNOMAN Monte Carlo was run to produce the response and to compare it to the original version of the code.

At first, only the scan at 386 nm was used in this χ^2 fit. The simulation was run at the correct energy to produce photons at this wavelength in order to fit the response to the calibration data. This wavelength was chosen since it is close to the peak of the PMT response. Figure 4.6 shows the distribution of wavelengths of the Čerenkov photons that cause successful PMT hits. The generation of Čerenkov photons is proportional to the inverse square of the wavelength but this distribution is then distorted by absorption in the AV and further altered by the wavelength-dependent detection efficiency (the quantum efficiency) of the PMTs. The most probable wavelength to cause a PMT hit is between 350-400 nm. Using 386 nm to tune the response was therefore a good approximation, but the response at 365 and 420 nm in particular was also important.

The response produced using the MCE parameters found in the one-wavelength fit was tested at these three central wavelengths and was found to be an adequate match to the data, but not a significant improvement over the previous form. To optimise the response across the range of most probable wavelengths, the method was extended to allow a simultaneous fit at all six wavelengths for which calibration data had been taken. This involved introducing a scaling to allow for the relative normalisations of each wavelength. In the single wavelength fit, both the data and the simulated response were normalised to 1.0 at normal incidence, so that only the shape of the response was fit for. The calibration data was provided in this form, showing only the shape of the response, not the overall collection efficiency. The simulation includes several wavelength-dependent factors, including the photocathode response, so in order to fit all six wavelengths simultaneously, the calibration data at each wavelength was scaled by the quantum efficiency of the PMT (the probability that an absorbed photon causes a photoelectron) at that wavelength, to give the correct normalisation.

A weighting was included in the fit, obtained from the information in figure 4.6, in order to account for the wavelength-dependent response of the PMTs. When calculat-

Wavelength Distribution for Salt N16 in the Center
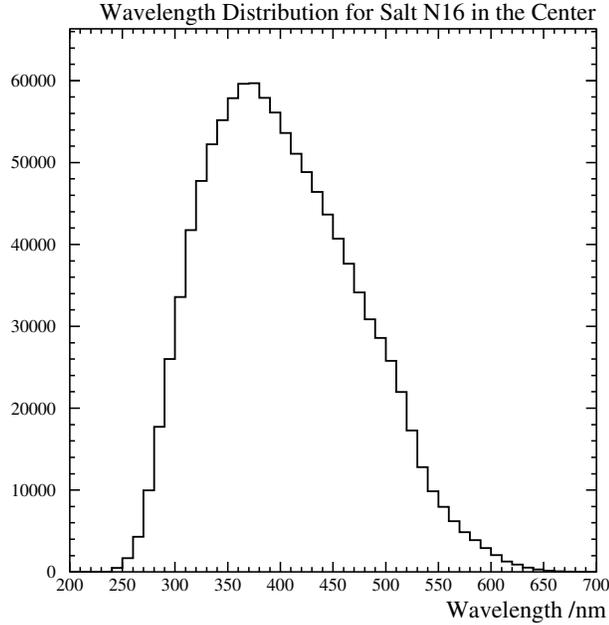
Figure 4.6: Wavelength dependence (/nm) of the PMT response, generated using ^{16}N events in the centre of the detector in the salt phase. Figure courtesy of M. Dunford.

ing the χ^2 for the global fit, instead of weighting the χ^2s of the fits at each individual wavelength equally, each wavelength was weighted by the relative likelihood of a hit being caused by a photon at that wavelength. This resulted in a greater weighting for the more probable wavelengths and a very small weighting for the data at 620 nm, for example, where the probability of a photon causing a PMT hit is very low. This method was found to produce very accurate and reliable results for the angular response across all the most probable wavelengths.

As a further extension, the inclusion of a wavelength-dependence to the fit was investigated. In addition to the variance of the response with hit position incorporated by the empirical MCE function, this method allowed the response to vary as a function of the wavelength of the incident photon as well. Although this method improved the fit at certain wavelengths, the improvement was small and was not consistent across all the optics scans. It was also found [67] that the wavelength dependence caused the overall normalisation to vary from scan to scan, with no physical basis for the change, affecting the energy scale in a time dependent fashion. Therefore, it was decided that the non-wavelength dependent fit was preferable and it was adopted as the final method for this analysis.

4.3.2 Angular Response for New Concentrators (D$_2$O Phase)

Only one full laserball scan exists from the D$_2$O phase in SNO. Since this is the first set of calibration data, it essentially corresponds to when the concentrators were new. Although there was a known drop in light collection efficiency during the D$_2$O phase that may have been due to the concentrators, the calibration data was used to represent the average response across the livetime of the phase, since no further data was available in order to model any time-dependent degradation.

Using the six-wavelength fit described in section 4.3.1, the five free parameters in the MCE were tuned to match the data at all wavelengths. The resulting alteration to the MCE is shown in figure 4.7. This plots the value of the MCE for each set of parameters as a function of *zpm*, the height at which the photon strikes the PMT. *zpm* is measured in centimetres from the origin, as marked on figure 4.1, so the equator of the PMT is at $zpm \sim 5.1$ cm and the centre of the front face of the PMT is at roughly 12.5 cm. Table 4.1 gives both the original values of the parameters used in the MCE function, when it was first incorporated to improve the modelling of the detector response, and the new values extracted from the fit.
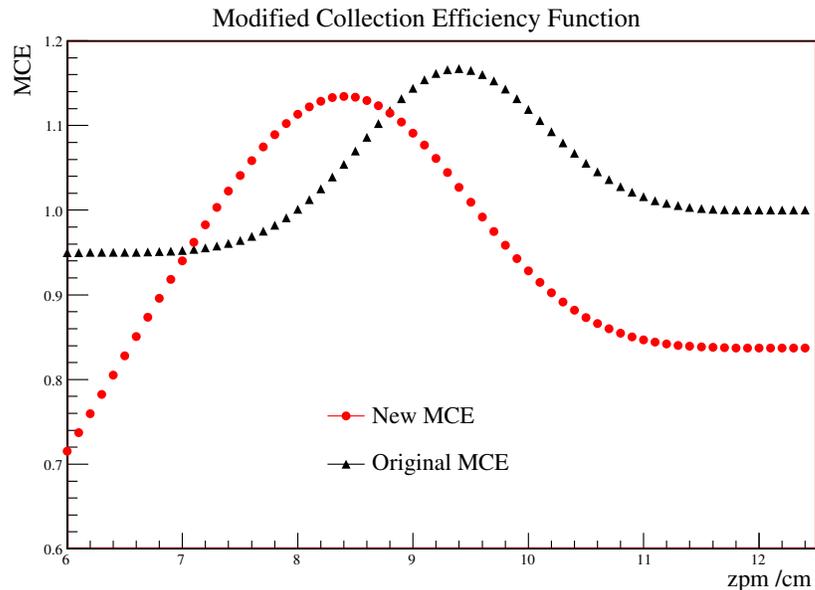
Figure 4.7: The modified collection efficiency as a function of the height at which a photon strikes the PMT, for the original and the new values of the parameters.

The resulting response for each set of parameters at each wavelength is shown in figure 4.8 and the values of the reduced χ^2 for the fit of each version of the simulation to the data are given in table 4.2. The errors on the data are the combined systematic and

	a_1	a_2	a_3	b_1	b_2
Old	0.48	0.43	0.95	-0.0175	-0.007
New	0.9230	0.4797	0.3939	-0.0095	-0.0007

Table 4.1: The values of the parameters used in the modified collection efficiency function for the D_2O phase angular response. 'Old' refers to the original values used in the function, before the improvements presented here.

statistical errors whereas the errors on the simulations are purely statistical. The plot at 620 nm is somewhat uneven. This is due to the rapid drop in the quantum efficiency of the PMT as the wavelength increases above 500 nm, which leads to poor statistics at these high wavelengths.

Wavelength /nm	337	365	386	420	500	620
χ^2 for old MCE	60.5	608	2.59	701	207	2.65
χ^2 for new MCE	1.02	2.19	0.89	23.4	15.1	99.8

Table 4.2: Reduced χ^2 values for the fit of the simulated angular response curve to calibration data for each wavelength, from 0–40°.

The fit is a significant improvement at all wavelengths except 620 nm, where the sensitivity of the PMT to photons is almost zero (see figure 4.6). The significance of the disagreement at 620 nm is nevertheless smaller than that seen at more probable wavelengths in the original version of the code. One possible explanation for the disagreement comes from the strong wavelength dependence of the attenuation length for Čerenkov photons in light water. At 620 nm, the attenuation length in H_2O is a few metres whereas at 500 nm it is roughly an order of magnitude greater. Small discrepancies in the photon paths could have a noticeable effect on the response at 620 nm, which would not be evident at the other wavelengths. Therefore, the discrepancy observed at this wavelength could be an indication that the path lengths calculated for photons in the light water are not absolutely correct.

4.3.3 Angular Response for Aged Concentrators (Salt Phase)

In the salt phase there were multiple laserball scans, so a time dependence of the angular response could be established. The fit from the D_2O phase was used as a starting point and the response was degraded to simulate the aging process observed in the data. Several of the laserball scans produced fairly similar shapes for the response, so rather than treating each scan individually, three were selected to represent the changes across the time span
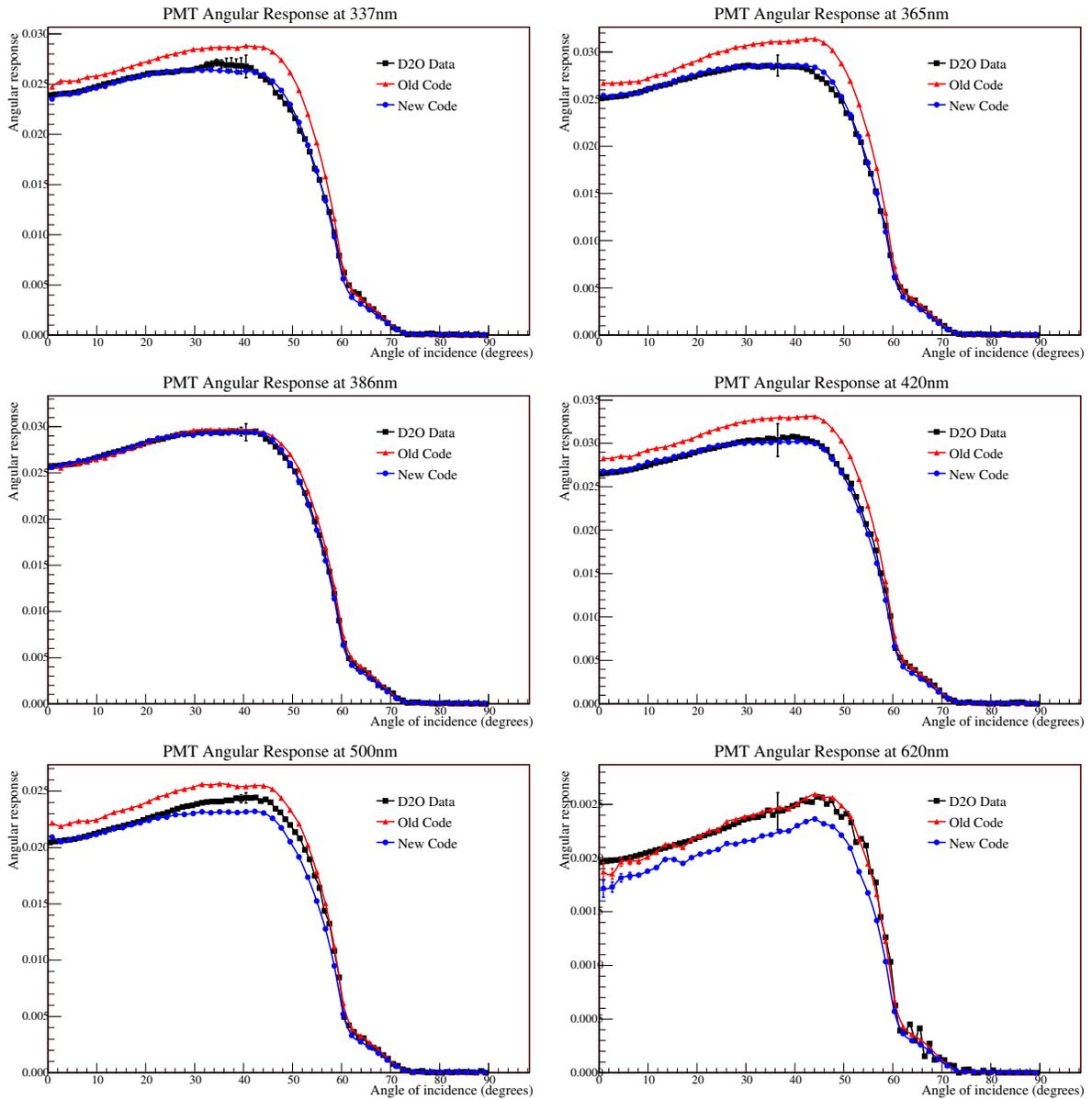
Figure 4.8: A comparison of calibration data to the Monte Carlo prediction for the angular response for the original and the optimised versions of the simulation code in the D_2O phase.

of the salt phase. The empirical nature of the MCE function meant that it was difficult to implement any continuous form of degradation. Instead, the salt phase was split into three discrete time windows and one set of calibration data was used to represent the optical response for each period. The time periods and the scan chosen to represent each one are shown in table 4.3.

Date range	25/07/01 - 17/01/02	17/01/02 - 3/03/03	3/03/03 - 29/08/03
Calibration scan	Sept 01	May 02	Apr 03

Table 4.3: The time periods selected for analysis of the angular response in the salt phase

To demonstrate the degradation of the response with time, the response at 386 nm, normalised to 1.0 at normal incidence, is shown for the D_2O phase and two of the salt phase scans in figure 4.9.

The six-wavelength fit was performed for each of the three salt phase scans, resulting in three sets of parameters for the modified collection efficiency. The resulting improvements in the match to calibration data were similar to those already presented for the D_2O phase. Figure 4.9 also shows an example of the improvement in the fit for a central wavelength in the April 2003 optics scan. For more detailed results of the improvements made to the simulated angular response in the salt phase, see appendix B.

4.3.4   Low Wavelength Discrepancies

The two energy fitters used for the LETA analysis (RSP and FTK) both used a set of look-up tables, known as *'Grey Disk'*, to determine the collection efficiency for a given wavelength and angle of incident light (see chapter 2). Therefore, it was important that the response modelled in these tables matched that predicted by the Monte Carlo simulation very closely, or a bias could be introduced between the estimated event energies and the true values. Since the energy fitters were applied to both data and Monte Carlo generated events, this bias would be reproduced in both and, therefore, should not affect the final flux analysis (which analyses the data in terms of what is predicted by Monte Carlo), but it still needs to be understood.

Such a bias was observed in the energy estimators, manifesting as a radial bias. Investigation into this effect by M. Dunford demonstrated that it was due to a discrepancy in the response at low wavelengths. For wavelengths above 337 nm, Grey Disk uses the actual calibration data to model the optical response and interpolates between the six wavelengths at which data was taken. Since the model in the simulation code was tuned to reproduce this response, the match between Grey Disk and the Monte Carlo prediction was very good. Below 337 nm, no calibration data exists to use in Grey Disk and, so, the shape of the 337 nm response was assumed and the normalisation was scaled down by the ratio of the PMT quantum efficiencies at the relevant wavelengths. The simulation code includes many wavelength-dependent factors (including, but not limited to, the photocathode response and the petal reflectivity) that can alter the optical response. The Monte Carlo prediction showed a significant change in the shape of the response at low
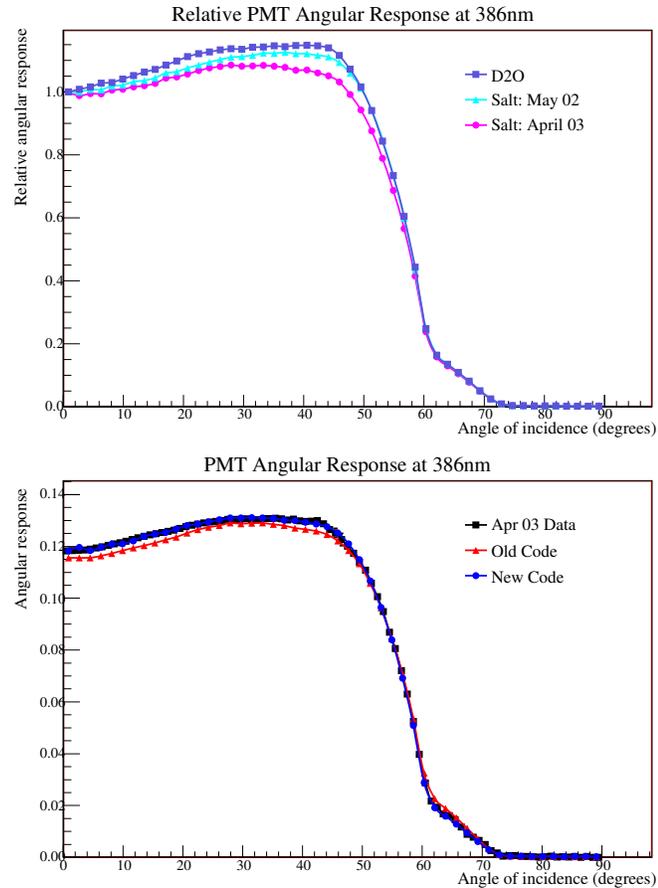
Figure 4.9: The effect of concentrator aging on the angular response. The upper figure shows the response in the D_2O phase and two of the salt phase scans to illustrate the degradation due to aging. The lower figure shows a comparison of the Monte Carlo prediction to calibration data for the original and the optimised versions of the simulation code, using the April 03 optics scan at 386 nm.

wavelengths, which was not reproduced in Grey Disk. Figure 4.10 illustrates the effect of this discrepancy on the detector response predicted by Grey Disk and by the full Monte Carlo simulation. A clear bias is observed between the Monte Carlo model and the Grey Disk model at high radii.

The possibility of a seven-wavelength fit was investigated, in which a response at 280 nm was constructed by scaling the 337 nm calibration data down by the ratio of the PMT quantum efficiencies to match what was used in Grey Disk. Including this in the fit to tune the MCE was intended to bring the simulation code closer into alignment with the prediction from Grey Disk. However, the weighting of this seventh wavelength was low in comparison to the central wavelengths (see figure 4.6) and so the fit was
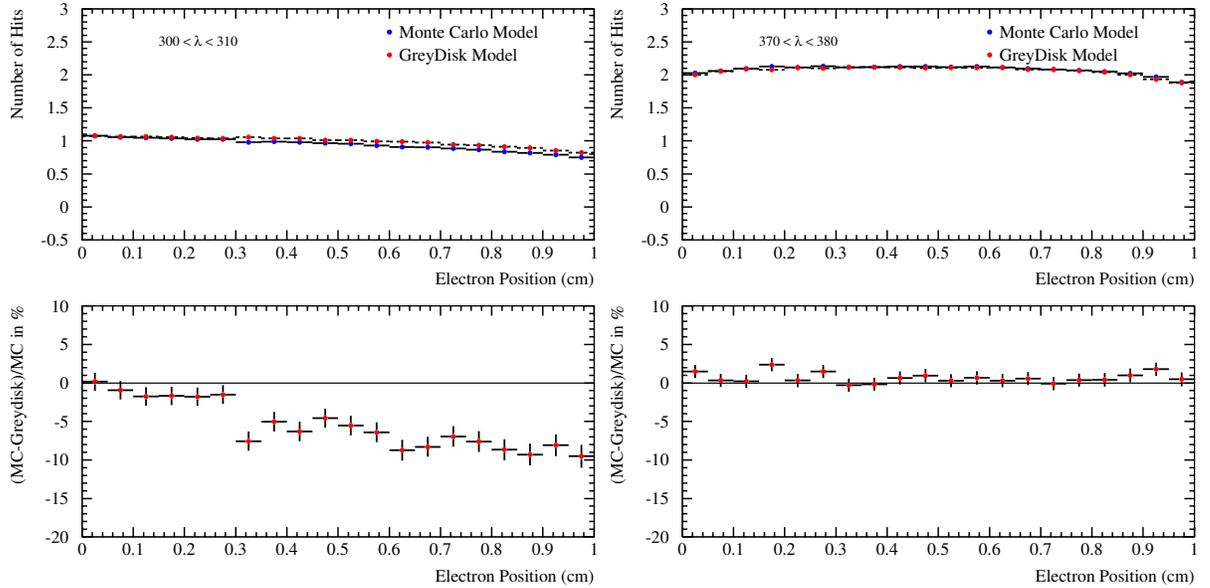
Figure 4.10: A comparison of the radial dependence of the energy scale for Grey Disk and the simulation code at 300-310 nm (left) and at 370-380 nm (right). Figures courtesy of M. Dunford.

still dominated by the response at 350-450 nm, resulting in a poor fit at 280 nm and no significant improvement in the bias.

There was no physical basis supporting the assumption that the response in Grey Disk was correct at this low wavelength. In fact, bench tests from several years previous had shown that the shape did continue to change as the wavelength decreased. Therefore, it was decided to use the Monte Carlo prediction for the shape in preference to the simple scaling used in Grey Disk. SNOMAN was used to generate look-up tables for Grey Disk at these low wavelengths, removing the discrepancy between the two models.

4.3.5   Impact of the Improvements in the Response

To understand the impact of the alterations presented in the sections above, some investigation was performed into the distribution of angles of incidence of photons causing successful PMT hits. Photons were generated uniformly and isotropically within the D_2O region of the detector and the angle of incidence of the photons that caused successful PMT hits was recorded. Figure 4.11 shows the distribution of these successful hits.

The peak of the distribution is between 20-40°, which is close to the peak in the PMT response. The improvement in the response in this range (as illustrated in figures 4.8 and 4.9) will therefore have an impact on the accuracy with which the overall detector response is modelled. The changes at low angles may be less significant to the overall
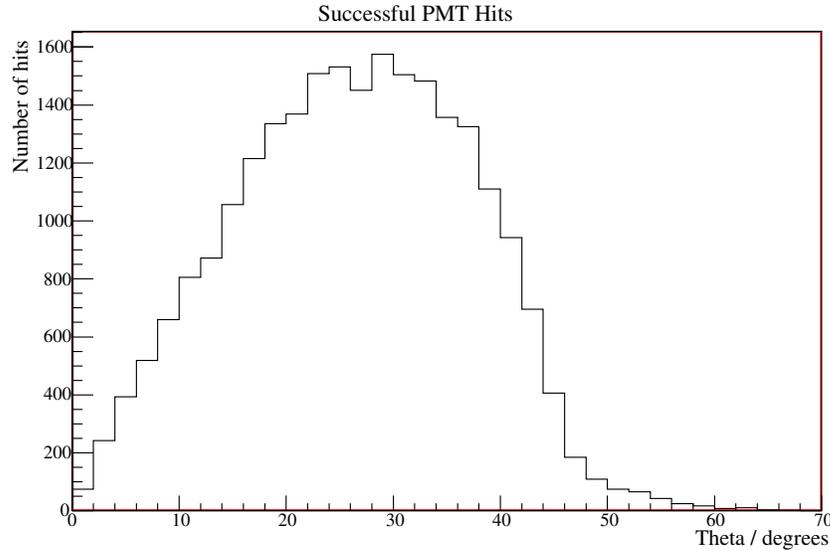
Figure 4.11: Angular distribution of photons incident on the photocathode that cause a successful PMT hit.

response, since fewer photons strike at these angles, but it was important to be very accurate in this region as the response at normal incidence was used to fix the overall collection efficiency.

Before the optimisations presented here, a radial bias existed between the detector response in data and the predictions from Monte Carlo. A mis-modelling of the angular response in the simulation was thought to have been the culprit, since a distortion in the shape of this response would affect the detection efficiency of the detector in a position-dependent fashion. This bias was a direct measure of the systematic uncertainty on the energy scale in the detector.

To demonstrate the improvement in the modelling of the detector response, M. Dunford investigated the effect of the changes on this bias. Using the RSP energy fitter, she compared the number of prompt PMT hits ($NWin$) in ^{16}N calibration data to Monte Carlo predictions for the original and the new angular response models. Figure 4.12 shows the changes in both the D_2O and salt phases.

The tuning of the angular response has reduced the bias in the energy scale of the detector dramatically in both phases of the experiment, as demonstrated by the reduction in the volume weighted mean of the difference between the data and Monte Carlo responses.
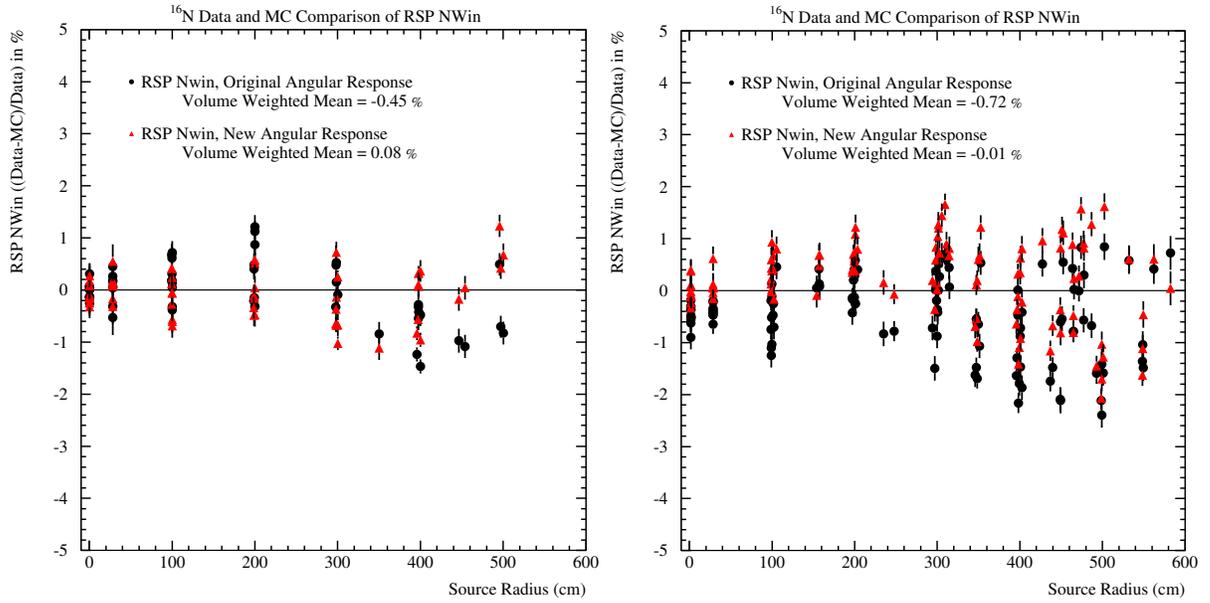
Figure 4.12: A comparison of the prompt PMT hits for ^{16}N data and Monte Carlo, before and after the angular response tuning in the D_2O (left) and salt (right) phases. Figures courtesy of M. Dunford.

4.4 Summary

Accurate estimation of event energies in SNO requires a full understanding and detailed modelling of the detector. A mismodelling of the optical response could introduce biases in the reconstructed energy, resulting in a distortion in the final extracted neutrino spectrum. This chapter has presented a full study of the angular response of the PMT and light concentrator setup, in order to fully understand the behaviour and to improve the accuracy with which it is modelled in the simulation.

Bench-top measurements of the angular response have demonstrated that discrepancies seen between previous measurements were consistent with PMT to PMT variations. It was also shown that the change in the response in the detector over time was consistent with the expected degradation of the concentrator petals due to immersion in water.

In addition, the model of the angular response in the simulation was optimised to better reproduce *in-situ* calibration data and to incorporate the effects of concentrator aging. As a result, the radial energy bias has been dramatically reduced. In previous analyses [12] a systematic uncertainty was incorporated to allow for this bias in the energy response. The optimisations presented here have rendered the inclusion of such an uncertainty no longer necessary. This has a direct impact on the accuracy with which the final neutrino spectrum can be measured, by reducing the uncertainty in the energy scale.

Chapter 5

Reconstruction

An important part of any attempt to understand the data taken by the SNO detector is the estimation of each individual event's location. This process is known as vertex reconstruction. A calculation is performed based on measured properties of the event, such as individual PMT position and timing, by an algorithm known as a fitter. As well as reconstructing the event's location, the fitter can also be used to determine other parameters such as the direction of the electron or γ-ray that caused the observed PMT hits. This chapter describes the calculation of the uncertainties associated with the reconstruction process and how they vary with both position in the detector and visible energy across the range of interest.

5.1   Reconstruction for LETA

Many of the background events in the SNO detector originate, by design, outside the main detector volume. The most efficient (and sometimes the only) way to identify these events is by their position. Defining a fiducial volume, as described in chapter 3, significantly reduces the tails of external events that reconstruct inside the analysis window. Any uncertainties in reconstructed event positions directly affect the number of events observed inside this volume. In particular, any biases that cause this number to change with energy have a direct impact on the energy dependence of the flux, distorting the extracted neutrino spectrum. The position distribution of events within the fiducial volume was also one of the elements used in the signal extraction procedure (discussed in chapter 8) to distinguish between different event types. A knowledge of the accuracy of the reconstruction process was therefore a vital ingredient to the final extraction of neutrino fluxes from the SNO data set.

5.1.1 The Fitters

The LETA analysis was run with two separate position fitters, as described in chapter 3, so the features of both must be investigated. The path fitter, FTP, is a maximum likelihood reconstruction algorithm that incorporates time and angular information to fit for event position and direction [52]. The FTU fitter reconstructs vertices based solely on the time at which inward looking PMTs fired [53].

The analysis also made use of two independent energy fitters: FTK and RSP. The most pertinent difference between these two fitters is that RSP uses only the PMT hits caused by prompt photons to calculate an event's energy whereas FTK uses scattered and reflected light as well [51].

These fitters were combined to form two independent analysis paths: the FTK/FTP path and the RSP/FTU path. This chapter discusses the position uncertainties associated with each analysis path; the uncertainties involved in the reconstruction of event energies is discussed elsewhere [47].

5.1.2 Analysis Method

The aim of this analysis is to determine the uncertainties associated with reconstruction. Since all the cuts intended to remove background events were applied to both data and Monte Carlo events, and since the Monte Carlo simulation was used to generate the PDFs used in the signal extraction procedure, it is any differences between the reconstruction of data and Monte Carlo events that could affect the final neutrino fluxes. Any biases in reconstruction that are well modelled in the Monte Carlo would be taken into account by the nature of the analysis. For example, if a bias in reconstruction caused all events to reconstruct 5 cm too high, but this was accurately modelled in the Monte Carlo, then it would be equivalent to shifting the entire fiducial volume down by 5 cm in both data and the Monte Carlo simulation.

Since both data and Monte Carlo events were required for this analysis, information from detector calibration sources was used. A direct comparison was made of the action of the fitters on source data to the model predictions for that source and the difference was taken as an estimate of the systematic uncertainty.

Reconstruction uncertainties for both electron and neutron-like events were investigated since the results were applied to both types of PDFs in the final signal extraction. The uncertainties were evaluated separately in each phase since the detector response changed over time. The ^{16}N calibration source was used to model the electron-like events in both phases. ^{16}N is essentially a monoenergetic γ source, whereas ^{8}Li produces electrons with a spectrum comparable to the ^{8}B spectrum and, so, the ^{8}Li source was used

to investigate any energy dependencies for these events. In the D_2O phase, the capture of a neutron on deuterium resulted in a single γ of roughly 6.25 MeV, which is well modelled by the ^{16}N decay. However, in the salt phase the neutrons captured predominantly on chlorine, which has a more complicated decay scheme, emitting multiple gammas at higher energies. This was modelled using the ^{252}Cf source, which produces a neutron burst as described in section 2.3.2.

To ensure a clean data sample of calibration events without any anomalous backgrounds, various cuts were applied to the data. For tagged sources, a cut exists to select only events for which the source trigger fired (see section 2.3.2). Further analysis cuts were applied to remove both instrumental and physics backgrounds (see chapter 3 for a description of these cuts). The source trigger and instrumental backgrounds are not simulated in the Monte Carlo and, so, these cuts were applied to data events only. The high level cuts, which predominantly remove physics backgrounds, were applied to Monte Carlo events as well as to the data. These cuts remove a small proportion of signal events as well as backgrounds, which can alter the distributions of the events passing the cuts. Such effects need to be reproduced in the simulation so that the comparison to data is unbiased. No fiducial volume cut was applied in this analysis, since its application would bias the reconstructed position distributions near the edge of the volume.

For events that passed all these cuts, the difference in the reconstructed position of each event from the known source position was histogrammed in the x, y and z planes and a function was fit to the resulting distributions. The function used in the fit depended on the source in question; this is discussed in more detail below. In all cases, the function had two parameters that were allowed to vary in the fit, representing the mean reconstruction error (the average offset of the reconstructed positions from the source position) and the resolution of reconstruction (the width of the distribution). From previous studies, the detector resolution was expected to be on the order of 20–30 cm. Therefore, a criterion was applied to select good fits, requiring a fitted resolution between 5 and 70 cm. This removed low statistics runs from the analysis, which tended to return unphysical values for the fit parameters. The reconstruction offset and resolution were then compared for data and Monte Carlo events on a run-by-run basis.

Since reconstruction effects were studied along the three detector axes individually, runs were split into two groups: single and dual-axis runs (defined in section 2.3.2). Single axis runs were taken along the z-axis only, at a fixed (x,y) position. These runs have a more precisely measured source position in the data, accurate to roughly 2 cm. Dual axis runs cover the full detector volume but the stated source positions are accurate to roughly 5 cm in the inner parts of the detector and 10 cm in the top regions of the vessel [68].

5.2 Source Distribution Functions

The calibration sources generate events with a known position distribution. The source distribution function is the distribution of positions around the source at which Čerenkov light is first emitted. For each source, this function was convolved with a hypothetical function representing the detector resolution and the result was fit to the extracted position distributions of data or Monte Carlo events. The resolution function had a mean, μ, and width, σ, which were allowed to vary in the fit. These parameters represented the reconstruction offset and resolution, respectively.

The detector response function used was a simple Gaussian centred on the point at which the light was produced. Although Monte Carlo studies of reconstructed electron position distributions showed that a better fit was obtained by using a Gaussian with an exponential tail [69], this would add extra complexity to the fit by introducing two further fit parameters. Since the purpose of this study was to estimate the uncertainties in reconstruction characteristics rather than to construct actual detector response functions, this added complexity was deemed unnecessary. Instead, the Gaussian was fit over a restricted range ($\pm 50\,\text{cm}$) around the peak of the distribution so as not to overweight the effect of the non-Gaussian tails.

The sources used for reconstruction were the ^{16}N, ^{8}Li and ^{252}Cf sources. The ^{16}N source is a γ-emitter and, so, the detected event is the first vertex at which a photon Compton scatters an electron with a high enough energy for the electron to be above the Čerenkov threshold. The distribution was extracted from the Monte Carlo simulation by generating ^{16}N events in the centre of the detector and histogramming the x, y and z positions at which Čerenkov light was first emitted. Some further complications were encountered with this distribution, which resulted in small alterations to this method; these are discussed below. Figure 5.1 shows a typical fit to ^{16}N data using this function convolved with a Gaussian, as described above.

^{8}Li is an electron source, with the events produced on an approximately spherical shell around the source. An analytical model for this behaviour exists, but the use of the ^{8}Li source in this analysis did not require the full source distribution function, as will be seen below.

^{252}Cf is a neutron source, as described in section 2.3.2. The neutrons produced tend to wander a finite distance before capturing and therefore the source distribution function is more complex to model accurately. However, a simple Gaussian fit in a narrow region around the peak of the reconstructed position distribution provided a good approximation and allowed the mean reconstruction offset to be determined with sufficient accuracy.
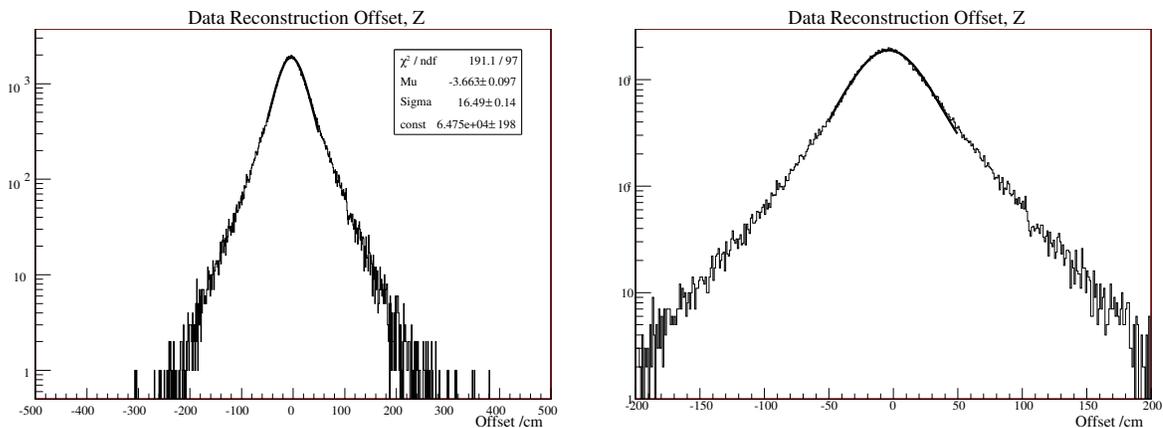
Figure 5.1: Fit of the ^{16}N source distribution function convolved with a Gaussian to the (reconstructed - source) position along the z-axis, for ^{16}N run 14670 in the D$_2$O phase. The full distribution (left) and a narrower range (right) are shown.

5.2.1 Reconstruction Using ^{16}N

A first look at reconstruction using the ^{16}N calibration source in the salt phase yielded some unexpected results. Strong time dependencies were observed in both the reconstruction resolution and offset along each of the three axes. Figure 5.2 shows the trends observed along the z-axis across the full time span of the salt phase. A zig-zag trend is clearly illustrated, as is an overall degradation of the response, indicated by an increase in the discrepancy of the reconstruction offset from zero. Some degree of degradation was expected due to PMT loss, concentrator degradation, attenuation drifts and other effects, which is not a problem as long as these effects are accurately modelled in the simulation. Although the trends were reproduced in the Monte Carlo, they were far stronger than would be expected and the zig-zag shapes could not be explained by the above processes. Their existence implied an incomplete understanding of some part of the reconstruction process, which warranted investigation. These trends were only seen when analysing with FTK/FTP; the RSP/FTU analysis yielded results that were satisfyingly constant over time, excepting for a small, fairly constant degradation, which could be explained by an overall drop in the number of working tubes (see section 5.2.2). As a further comparison, the results for both fitters in the D$_2$O phase yielded results that were constant over time.

In addition to this time dependence, the reconstruction resolution for FTU was observed to degrade by approximately 2 cm between the D$_2$O and salt phases, but only in the simulation. The data resolution stayed roughly constant between the two phases. For FTP, the data resolution was well reproduced by the simulation but was several centimetres broader in salt than in the D$_2$O phase.

Figure 5.2: Time dependence of the mean reconstruction offset (left) and the resolution (right) along the z-axis for the FTK/FTP fitters in the salt phase, using ^{16}N data.

5.2.2 Resolving the Problem

One time dependent factor in SNO was the number of working PMTs in the detector. This could account for a zig-zag trend as tubes were switched off in the data stream due to problems with the electronics, which were fixed periodically. Figure 5.3 shows the number of working tubes across the span of the salt phase. The overall degradation that was observed, particularly in the RSP/FTU analysis, can be explained by the overall reduction in the number of working tubes across the duration of the phase. Although a zig-zag trend could be inferred, it did not explain the time dependent features observed with FTK/FTP. In particular, the runs with fewer working PMTs actually corresponded to runs that appeared to have better position resolution, contrary to expectation.
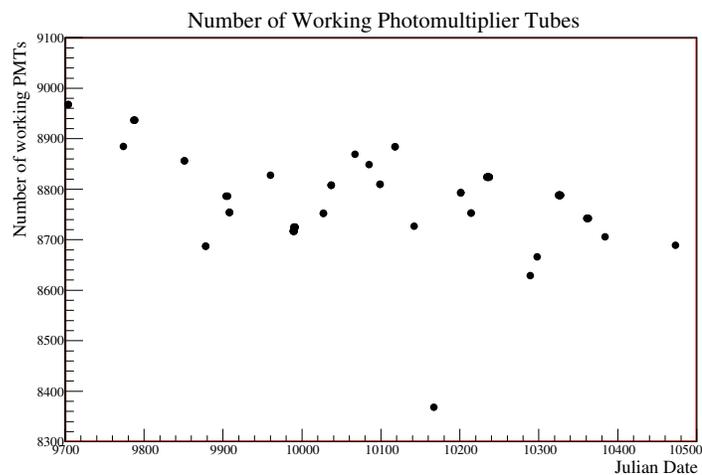
Figure 5.3: The number of working PMTs during the salt phase

The specific cuts applied to the simulation when extracting the ^{16}N source distribution function could have an impact on the shape of the resulting function, which would, in turn, affect the extracted values for the offset and resolution. In particular, the high level cuts used in the LETA analysis were known to remove some fraction of signal events and to be affected by the source geometry to some extent in both data and Monte Carlo simulations. When extracting the ^{16}N source distribution function as previously described, only the basic energy cuts were applied to the simulation to select events. To take these factors into account, the event selection was extended to include the full set of high level cuts. In addition, to allow for any time variations the distribution was extracted individually for each ^{16}N run, using the specific run conditions such as the correct number and distribution of working PMTs for that run. The Monte Carlo events used to extract the source distribution function were generated at the stated source position for each run instead of at the centre, as before, to also incorporate any variations in response across the detector volume.

This more accurate method for extracting the ^{16}N source distribution function was shown to completely remove the time dependent effects discussed above, as can be seen by a comparison of figure 5.4, produced using this method, with figure 5.2. A weak increase can be seen in the resolution as the detector response degraded over time, but this is well reproduced by the simulation. The new method also brought the salt results much more into line with the D$_2$O phase, removing the offsets between the two phases that had been seen in the resolution in particular.
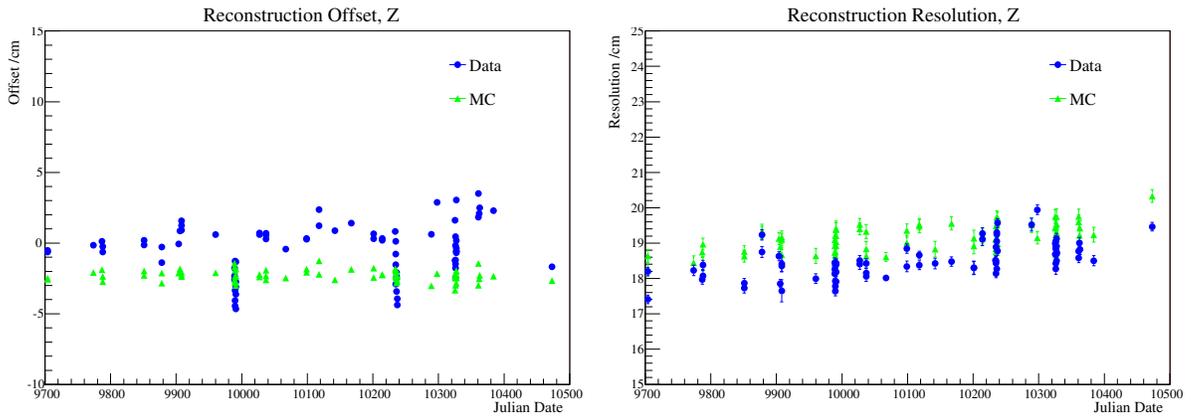
Figure 5.4: Time dependence of the mean reconstruction offset (left) and the resolution (right) along the z-axis for the FTK/FTP fitters in the salt phase, using ^{16}N data.

The difference was found to be due to the FTK energy uncertainty cuts (discussed further in sections 6.5 and 6.7), which were known to be sensitive to specific time-dependent run conditions such as the position of the source geometry and the number

of working PMTs and also to specific regions of asymmetry in the detector geometry, such as the neck. By not accounting for these factors in the original modelling of the ^{16}N source distribution function, a bias had been introduced into the function that affected the extracted values for the reconstruction parameters. This affected both data and Monte Carlo events similarly, since the same source distribution function was used for each. By fully accounting for all the time-dependent variations in run conditions in the extraction of the source distribution function, the bias was removed. Therefore, this method was used to generate the source distribution function for the final reconstruction analysis described in the remainder of this chapter.

5.3 Vertex Shift

The vertex shift is the average offset between the source position and the reconstructed positions of events. This was found using the ^{16}N source by taking a set of runs in each phase and extracting the mean and resolution for each run independently for each fitter. Several previous analyses, including a study of the energy scale [47], showed significant trends with z-position within the detector. For this reason, the vertex shift was not studied in terms of radial position, as it was in previous reconstruction analyses, but instead was determined for each of the three separate axes: x, y and z.

The vertex shift can vary with position in the detector. However, this effect was taken into account in the fiducial volume measurements described in section 5.6. Therefore, for the vertex shift calculation the analysis was limited to using central runs so as not to double count this effect. Only runs within 25 cm of the centre of the detector along each axis were used for that axis. As discussed above, the required result was a measure of the difference observed between the results for data and Monte Carlo events and, so, the difference between the extracted reconstruction offset for data and Monte Carlo events was found for each of these runs. The final vertex shift was then calculated as the weighted mean of these differences, along with the associated uncertainty.

The vertex shift was applied in the signal extraction as a systematic uncertainty, to take into account possible differences between reconstructed positions in the data and the simulation. Event positions in the Monte Carlo PDFs were shifted by a fixed distance to determine the effect on the extracted neutrino fluxes. The size of this systematic uncertainty was estimated using the observed differences between the simulation and the data in this analysis. The vertex shift calculated above was combined with the uncertainty on that value to produce a two-sided bound on the magnitude of this effect. In all cases, the reconstruction offset in the data was larger than in the simulation, resulting in a positive vertex shift. Therefore, the positive one sigma bound was estimated as the vertex shift

plus the associated uncertainty and the negative one sigma bound was just the negative uncertainty. The values for each phase and fitter are given in table 5.1 below.

Vertex Shift	D$_2$O		Salt	
/cm	FTU	FTP	FTU	FTP
X shift	+1.12 / -0.11	+1.15 / -0.13	+0.74 / -0.07	+0.62 / -0.07
Y shift	+2.59 / -0.15	+2.87 / -0.17	+2.10 / -0.08	+2.29 / -0.09
Z shift	+2.30 / -0.13	+2.58 / -0.15	+2.96 / -0.15	+3.11 / -0.16

Table 5.1: Vertex shifts in cm for the FTU and FTP fitters in the D$_2$O and salt phases.

5.4 Vertex Resolution

The vertex resolution is the width of the distribution of reconstructed positions around the true source position. This was found using a similar method to that employed for the vertex shifts, but considering the fitted resolution, σ, instead of the offset. Any position dependence in the resolution must be taken into account here, so all runs were used for this calculation, not just the central runs.

The application of the vertex resolution systematic uncertainty in the extraction of final neutrino flux uncertainties involved smearing the simulated position distribution by a Gaussian with a finite width, representing the difference between the measured detector resolution for data and Monte Carlo events. Since the Monte Carlo distribution cannot be '*unsmeared*', the resulting difference in extracted neutrino fluxes was applied as a symmetric uncertainty. The required result from the reconstruction analysis was therefore the width of the Gaussian by which the Monte Carlo distribution should be smeared to reproduce the data. This width is given by:

$$\sigma_{s,i}^2 = \sigma_{D,i}^2 - \sigma_{MC,i}^2 \tag{5.1}$$

where $\sigma_{s,i}$ is the width of the Gaussian by which the Monte Carlo distribution should be smeared to reproduce the data, $\sigma_{D,i}$ is the measured resolution in the data and $\sigma_{MC,i}$ is the resolution for Monte Carlo events along axis i of the detector ($i = x$, y and z).

In previous reconstruction analyses, the data resolution was always broader than that for the Monte Carlo. Interestingly, for some runs in this analysis the Monte Carlo distribution actually appeared to be broader than the data. In most cases the differences were so tiny that this effect was considered to be non-physical and purely a feature of the fact that the source Monte Carlo now so closely resembles the data that statistics

can lead the difference to fall one side or the other of zero. Since the application of this uncertainty to the neutrino flux extraction involved the assumption that the effect was symmetric anyway, the magnitude of the difference between the squared resolutions was used when calculating σ_s according to equation 5.1.

The value of σ_s^2 was calculated along each axis for each ^{16}N run individually, along with an associated uncertainty, and the result was averaged across all runs. Conservatively, the value quoted for use in the final neutrino flux extraction was the resulting value of σ_s plus the associated uncertainty. These values are given in table 5.2 below, along with the average extracted values for the data resolution.

Vertex Resolution	D_2O		Salt	
/cm	FTU	FTP	FTU	FTP
X data resolution	23.4	19.8	24.1	20.6
Y data resolution	23.6	19.9	24.1	20.7
Z data resolution	21.2	18.3	21.8	18.6
X smearing	4.1	3.3	4.9	3.1
Y smearing	2.6	2.2	5.3	3.4
Z smearing	3.8	1.5	8.0	5.3

Table 5.2: Vertex resolution in the data and the width of the Gaussian by which the Monte Carlo distribution should be smeared to model the (data - Monte Carlo) differences. Results are quoted for the FTU and FTP fitters in the D_2O and salt phases.

5.5 Consistency Checks on Vertex Reconstruction

Both the vertex shift and the resolution required a few consistency checks to ensure that the results from the ^{16}N source could reliably be applied to the neutrino data. The first of these was a check to ensure that the results remained constant across the time span of the data. The time dependence in the salt phase was explored in section 5.2.2. Figure 5.5 shows the time dependence of both the reconstruction offset and resolution in the D_2O phase for the FTK/FTP analysis. There are clearly no changes with time that need to be taken into account and the shift observed in the offset between data and Monte Carlo events is consistent with that found in section 5.3.

The ^{16}N source is a high rate source, generating events on a much more frequent basis than the expected rate of neutrino events within the detector. Therefore, a further check was required to investigate possible rate effects from the source. A few special runs were taken with the ^{16}N source at a lower rate and the analysis was repeated using these runs to compare to the high rate results. Figure 5.6 shows the reconstruction offset and
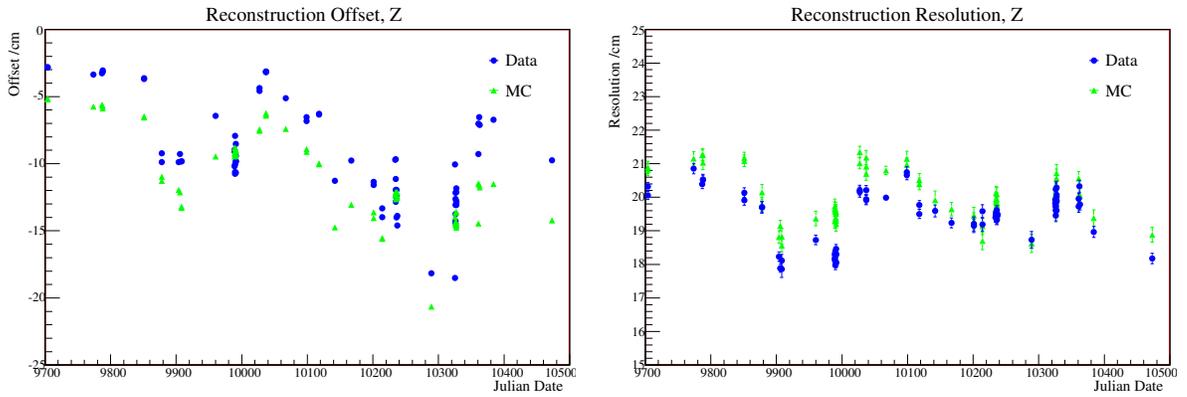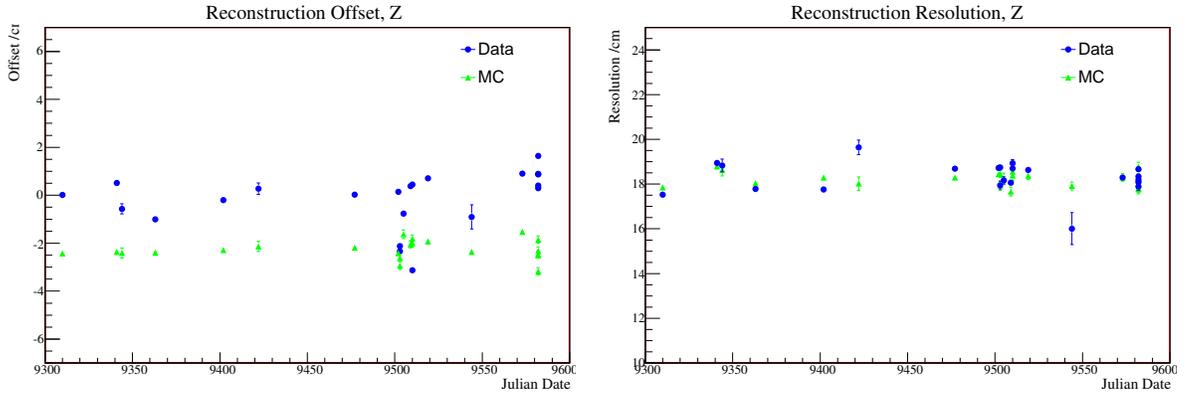
Figure 5.5: Time dependence of the mean reconstruction offset (left) and the resolution (right) along the z-axis for the FTK/FTP fitters in the D_2O phase, using ^{16}N data.

resolution as a function of position within the detector for these low rate runs for the FTK/FTP analysis in the D_2O phase. The statistics for the low rate runs are naturally lower, so the uncertainties are correspondingly larger. Table 5.3 shows the vertex shift between data and Monte Carlo and the data resolution for both the low rate and high rate runs, to illustrate the comparison. The numbers shown are for the FTK/FTP analysis in the D_2O phase only, but those for the RSP/FTU analysis and the salt phase were in similarly good agreement.
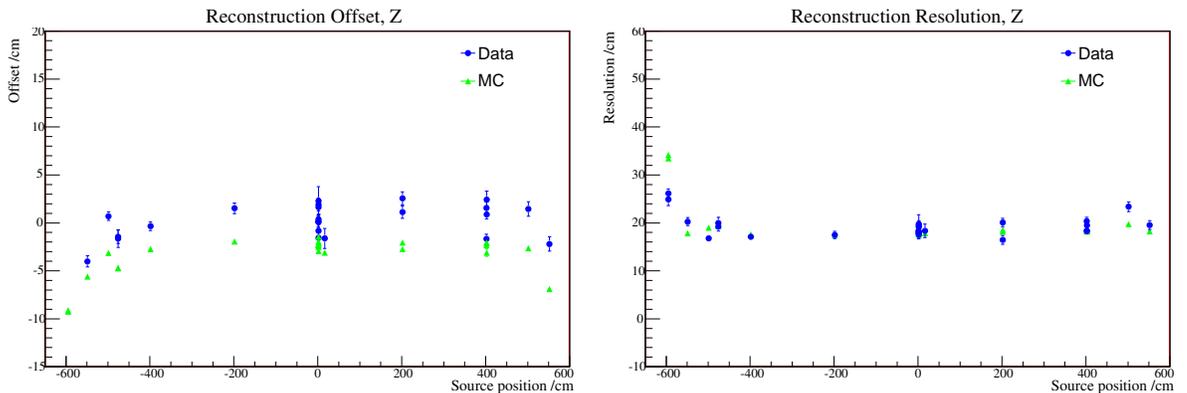
Figure 5.6: Low rate check of the mean reconstruction offset (left) and the resolution (right) along the z-axis for the FTK/FTP fitters in the D_2O phase, using ^{16}N data.

Other source effects can impact the results but previous work has been done to confirm that the effects are negligible [69]. This involved repeating the above analyses using the ^8Li source to demonstrate that the results were consistent for both calibration sources.

Reconstruction Parameter	Low Rate	High Rate
X vertex shift /cm	1.2 ± 0.3	1.0 ± 0.1
Y vertex shift /cm	2.0 ± 0.5	2.7 ± 0.2
Z vertex shift /cm	2.3 ± 0.4	2.4 ± 0.2
X data resolution /cm	20.6 ± 0.6	19.8 ± 0.1
Y data resolution /cm	21.0 ± 0.7	19.9 ± 0.1
Z data resolution /cm	18.9 ± 0.4	18.3 ± 0.2

Table 5.3: Low rate checks of the vertex offset and resolution for the FTK/FTP fitters in the D_2O phase.

5.6 Fiducial Volume

As well as a systematic offset in the reconstructed positions of events, the accuracy of vertex reconstruction can display a position dependence, which could differ between data and the Monte Carlo simulation. This was an important feature to investigate since a variation of this form could lead to a difference in the proportion of signal events reconstructing inside the fiducial volume between data and the Monte Carlo, which would bias the results from signal extraction. In the past, this possibility was modelled as a radial scaling of events, so that the reconstructed positions of events generated at the same position were related as follows:

$$R_{data} = \alpha \times R_{MC} \tag{5.2}$$

so that a value of α of less than 1.0 would mean that events in the data were being systematically reconstructed closer to the centre of the detector than those in the simulation.

Again, due to the strong trends seen in various parameters along the z-axis in the detector, this scaling was not studied radially but instead was split into three axial scalings. To remove any correlations between the axes, a cut was placed so that only runs within 50 cm of the axis in question were considered.

One of the major problems inherent in using calibration source data for reconstruction measurements is that there is always some uncertainty in the stated position of the source. Along the z-axis this can usually be quoted to within a couple of centimetres, but further out along the lateral axes it can sometimes be uncertain to 5 or even 10 cm. This can have large effects on the results of reconstruction studies. One thing that can be determined more accurately than the absolute source position is how far the source has moved between different runs in the same scan, *i.e.* the relative distances between source positions. For this reason, a new technique used in this analysis was to take the difference of every run from a central run *in the same scan*. This should remove any global offsets in the source position from the stated position and also removed the vertex shift

offsets discussed in section 5.3, thus ensuring that the two measurements were completely uncorrelated.

The method for this measurement was to select runs that were in a specific scan and choose a central run in that scan to be the point of comparison. For each run, the difference between the mean fitted positions for data and Monte Carlo events was found along a particular axis and that was then compared to the difference for the central run. The resulting value was plotted as a function of the reconstructed position for Monte Carlo events. This method could be repeated for several different scans and the results combined since any offsets between the scans should have been intrinsically removed.

The resulting graph should, by design, pass through the origin. Assuming a straight line dependence to the points, the following parameterisation was adopted:

$$X^i_{data} - X^i_{MC} = \beta \times X^i_{MC} \tag{5.3}$$

where X^i is the position on axis i ($i = x, y, z$) and β is the slope of the graph. This gives:

$$X^i_{data} = (1 + \beta) \times X^i_{MC} \tag{5.4}$$

such that $(1 + \beta)$ gives the radial scaling parameter, α, as described in equation 5.2.

In a comparison to previous studies [70], this method was shown to have significant benefits. In particular, a direct comparison between the analysis methods for the April 2003 ^{16}N scan showed that a clear distortion previously observed in the reconstruction offsets in data and Monte Carlo near the bottom of the detector was completely removed by the new method. This clearly illustrated that the cause of the distortion had been an error in the exact position of the source manipulator. Additionally, in the past each scan had to be treated individually and the values of radial scaling extracted for each scan varied widely. The results from the new method proved much more consistent across different scans and these results could be combined to further reduce the uncertainty by increasing the statistics of the measurement.

5.6.1   Electron-like Events

The ^{16}N source was used to investigate the scaling effects for electron-like events within the detector. This source was deployed in a wide range of positions, sampling most of the detector volume. For this measurement especially, it was very important to measure the effect as close to the edge of the fiducial volume as possible because any differences in data and Monte Carlo reconstruction near the edge would most directly affect the total flux seen inside the volume.

For the z-axis, ^{16}N runs were taken at a wide range of positions from the very bottom of the detector right up into the neck. However, along the x and y-axes the positions were

more limited, in particular not extending past about ± 450 cm along each axis. This led to the question of whether any trends observed in the inner volume for x and y should be extrapolated out to the edge of the fiducial volume, or whether the result from the z-axis should be used to represent all three axes.

Figure 5.7 shows a comparison of the results for the three axes. From this figure it can clearly be seen that the z-axis is not typical of the three, so the decision was made to treat the three independently. Therefore, the x and y scalings were found by extrapolating the position dependence seen in the central region, where scans were taken, out to the edge of the volume.
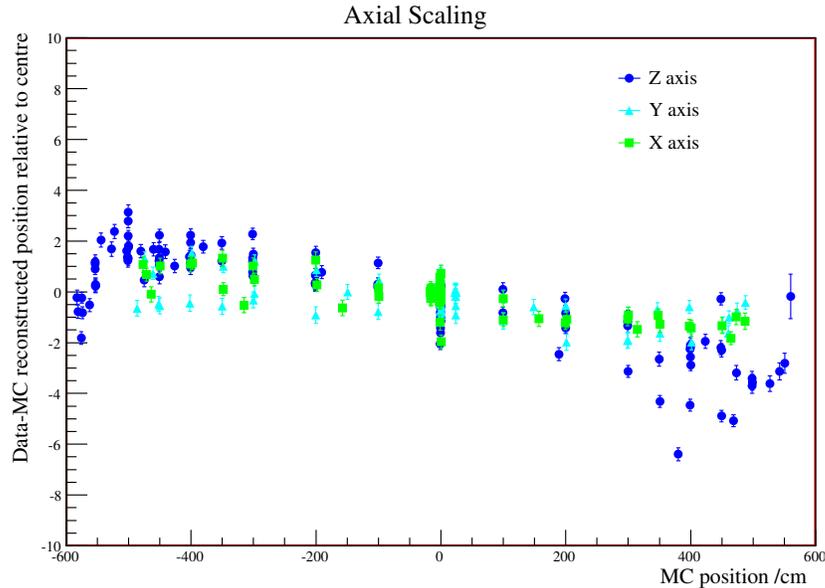
Figure 5.7: Comparison of axial scaling results for the three axes, using the FTP fitter in the salt phase.

A large number of ^{16}N runs were taken in both the D$_2$O and salt phases, as can be seen from the abundance of points in figure 5.7. This results in some spread in the possible value of the axial scalings. What was required from this analysis was a two-sided, one sigma bound on the values of the scalings that could be applied in the signal extraction as a systematic uncertainty on the simulated event positions. The two possible dangers were an overestimation of the uncertainties by incorporating every single ^{16}N run in the result or an overweighting of certain areas of the detector that might have a more extreme value of the scaling effect, but in which a relatively small number of events may occur. A solution to both of these concerns was found by splitting the detector into distinct volume elements along the axis in question. For example, along the z-axis the detector was split into horizontal slices, each with a height of 50 cm. All runs within the same

volume element were combined to give a weighted average for that bin. The result in each bin was then weighted by the volume represented by that bin.

This method was used to obtain results for each of the three axes in each of the two phases, with each fitter. A first order polynomial was fit to each of the resulting graphs. For the fits, the intercept was fixed to be zero since it was implicit in the method that any functional form should pass through the origin. The gradient from the fits gives a measure of the scaling parameter, β, from equations 5.3 - 5.4 above. Examples are shown in figures 5.8 and 5.9.
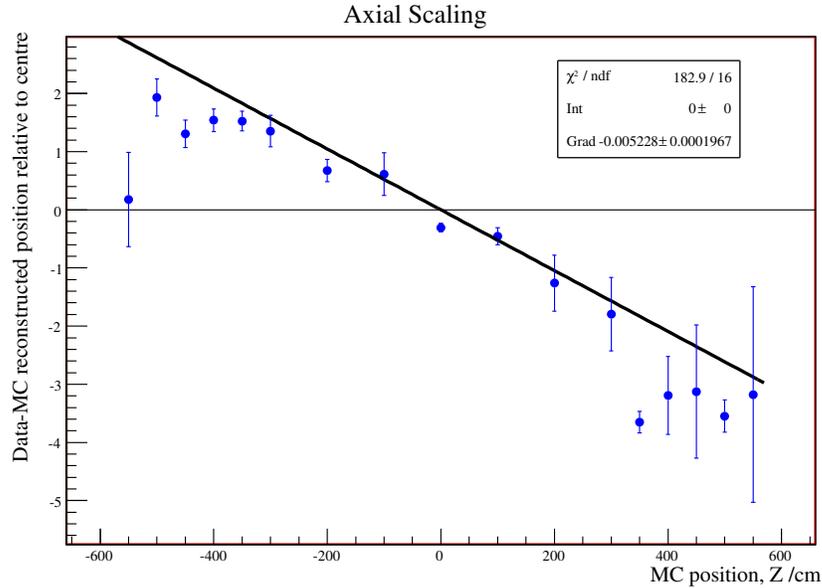
Figure 5.8: Axial scaling results for the Z axis, with volume weighted bins, using the FTP fitter in the salt phase.

Two methods were used for producing the final scaling values. For cases such as figure 5.8, the best-fit line is a good representation of the trend observed in the points. Although an apparent systematic scaling was observed along these axes, *i.e.* the gradient of the line was not consistent with zero, this effect was propagated in the signal extraction as a systematic uncertainty in the fitted position, not a correction. The value of this uncertainty for each axis was obtained by combining the gradient of the best-fit line with the fit uncertainty in that parameter. If the model was correct, the value of χ^2 per degree of freedom for the fit would be expected to be on the order of 1.0. A method used by the particle data group to allow for occasions in which a model is not a perfect fit to the data is to scale the uncertainty on the fit parameters by the square root of the χ^2 per degree of freedom. This is equivalent to scaling the errors on each point to a size at which the χ^2 per degree of freedom would have been 1.0, *i.e.* it is an extra uncertainty
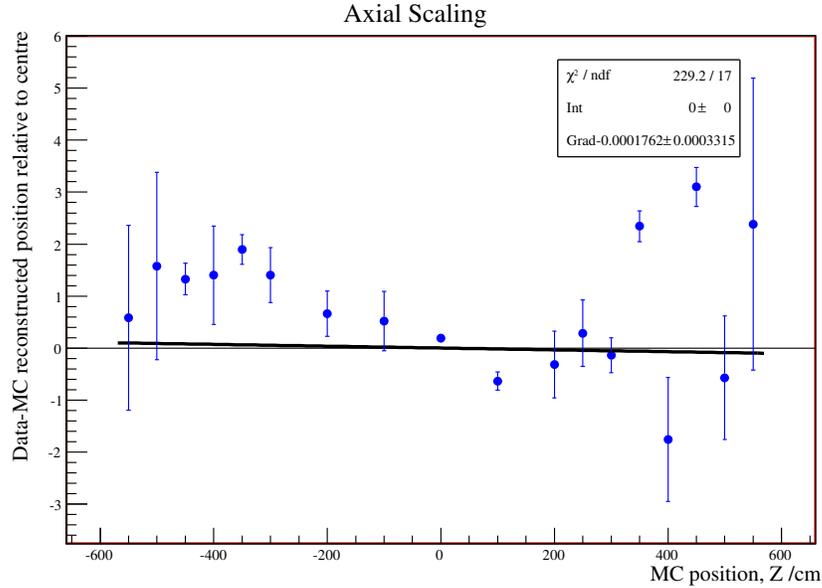
Figure 5.9: Axial scaling results for the Z axis, with volume weighted bins, using the FTP fitter in the D_2O phase.

to allow for the fact that the model in use is not a perfect representation of the data. The uncertainty on the slope was scaled in this way and then combined with the gradient to give the final value of the systematic uncertainty. The positive one sigma bound is quoted as just the positive uncertainty on the fitted gradient, since the slope is clearly in a negative direction. The negative one sigma bound was evaluated as the (negative) fitted gradient minus the scaled uncertainty, resulting in an asymmetric scaling uncertainty.

This method could not be used in all cases. In some instances, such as figure 5.9, the points were more evenly split above and below the axis and, so, the gradient of the best-fit line was much lower than might be expected and did not represent the spread in the points. For these cases, the best-fit line was discarded and the scaling was judged by eye and applied as a symmetric, two-sided uncertainty.

The full set of axial scalings found for electron-like events using ^{16}N is shown in table 5.4 below (the numbers quoted are values for the β scaling parameter in equations 5.3 - 5.4 above).

5.6.2 Diagonal Axis Verification

Since only the runs along each of the three axes were used to calculate these numbers, that leaves the diagonal-axis runs available to use as a confirmation of the results. Fairly few runs were taken along these axes in the D_2O phase, but a reasonable number were

Axial scalings	D$_2$O		Salt	
/%	FTU	FTP	FTU	FTP
X scaling	+0.07 / -0.46	+0.08 / -0.57	+0.04 / -0.26	+0.04 / -0.34
Y scaling	+0.09 / -0.41	+0.10 / -0.52	+0.04 / -0.23	+0.04 / -0.26
Z scaling	+0.50 / -0.50	+0.50 / -0.50	+0.50 / -0.50	+0.07 / -0.59

Table 5.4: Axial scalings for the FTU and FTP fitters in the D$_2$O and salt phases. Numbers are quoted in %.

taken during the salt phase. The method employed for this verification was very similar to that described above, with a rotation of axes so that either the y-z axis or the x-z axis was used instead of one of the three standard Cartesian axes.

A prediction can be made for what is expected by combining the results from the two Cartesian axes involved. The values in table 5.4 are given as fractional scalings of the positions along the standard Cartesian axes. These can therefore be combined in a straightforward way to produce a prediction for one of the diagonal axes. For example, rotating to the y-z axis, where ρ is the position on that axis and σ_ρ is the absolute uncertainty in that position, combination of errors gives:

$$\rho^2 = y^2 + z^2$$
$$\frac{\sigma_\rho}{\rho} = \frac{1}{\rho^2}\sqrt{\left[y^4 \left(\frac{\sigma_y}{y} \right)^2 + z^4 \left(\frac{\sigma_z}{z} \right)^2 \right]}$$

On the y-z axis, $y = z$ and, hence, $\rho = \sqrt{2} \times y$, so the above equation can be simplified to:

$$\frac{\sigma_\rho}{\rho} = \frac{1}{2}\sqrt{\left[\left(\frac{\sigma_y}{y} \right)^2 + \left(\frac{\sigma_z}{z} \right)^2 \right]} \tag{5.5}$$

Equation 5.5 can be used to predict the systematic scaling along any of the diagonal axes. Figure 5.10 shows the data for the negatively sloped y-z axis along with the best-fit to that data and the prediction calculated from equation 5.5 using the numbers in table 5.4. The prediction is a very good fit to the data, which strongly supports the use of the method described above for determining the magnitude of the systematic scaling along the individual detector axes. The very close agreement observed in this verification demonstrates that the method was both robust and accurate.
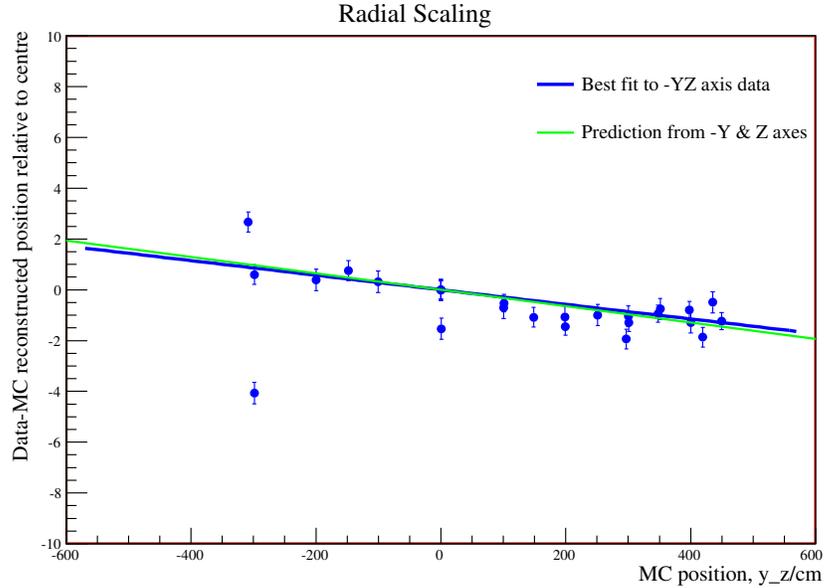
Figure 5.10: Axial scaling results for the negative y-z axis, using the FTP fitter in the salt phase. The data is compared to the prediction from the y and z axis results.

5.6.3   Neutron-like Events

Scaling effects within the detector could differ for neutrons and electrons and, so, the fiducial volume measurement was repeated for neutron-like events.

In the D_2O phase, neutrons captured on deuterium atoms and the subsequent decay of the excited state is well modelled by the ^{16}N source. However, in the salt phase the neutrons captured much more prominently on chlorine atoms, the excited state of which has a much more complicated decay scheme. Therefore the ^{252}Cf source was used as a neutron source in the salt phase.

The method used for the ^{252}Cf source was similar to that for the ^{16}N. Runs within a scan were selected and compared to a central run in the same scan, to remove any offsets. However, a few problems existed for the ^{252}Cf source, mostly due to the limited number of runs and the limited range of positions within the detector at which the source was deployed. It was not possible to split the runs into the three axes, as was done for ^{16}N, since most of the runs were either along the z-axis or diagonal axes so there would be no data left to use for the lateral axes. Instead, the runs were split into the z and x-z axes, which were the only two axes on which suitable runs (those in a scan with a central run) existed. Since these axes are not independent, correlations exist between the runs and so the points were combined for the analysis. The results of this analysis for the FTU and FTP fitters are shown in figures 5.11 and 5.12.
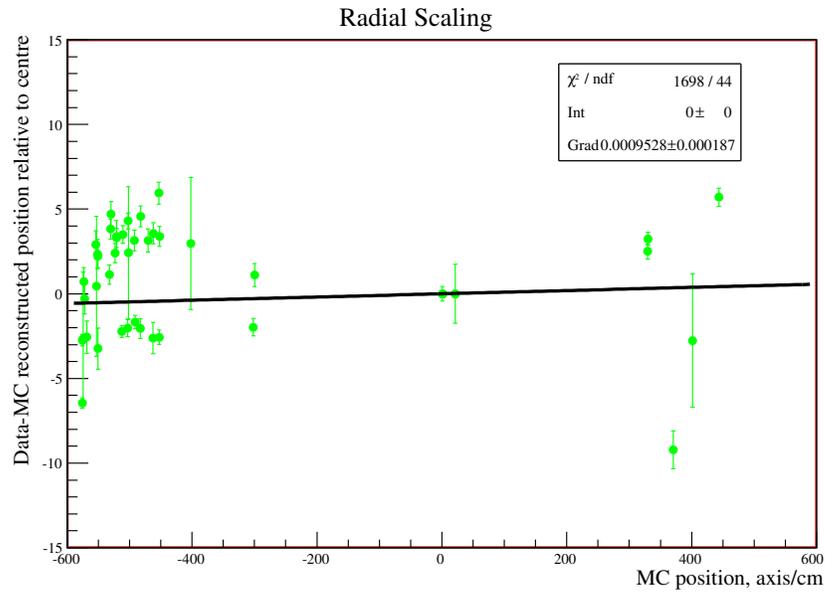
Figure 5.11: Axial scaling results for the ^{252}Cf source, using the FTU fitter in the salt phase.
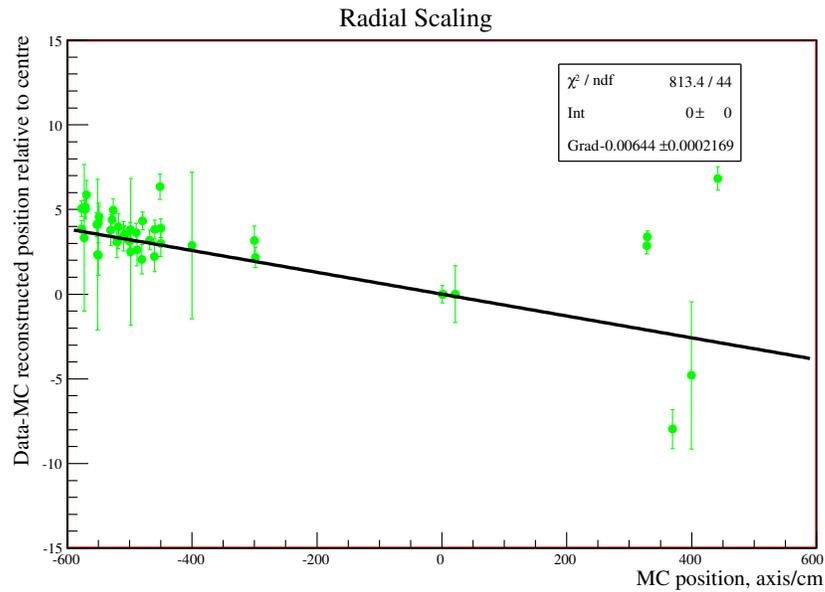
Figure 5.12: Axial scaling results for the ^{252}Cf source, using the FTP fitter in the salt phase.

In the case of FTU, the points are quite evenly spread above and below the axis, as was true for the ^{16}N source with FTU in the salt phase. Since all the ^{252}Cf runs were either

on the z-axis or the x-z axis, the z-scaling from the electron-like events is the pertinent one for comparison. For FTU in the salt phase, this scaling was $\pm 0.5\%$. By eye, it can be seen that this is a reasonable representation of the ^{252}Cf runs and, so, the electron scaling was taken to hold for neutrons as well.

For FTP, the points are distributed in a fairly clear line. The comparison from the ^{16}N source in the salt phase is a z-scaling of $+0.07/-0.59\%$. The best-fit line to the ^{252}Cf is actually slightly steeper than this, at -0.64%, but there are a few caveats to bear in mind. First, because of the smaller number of ^{252}Cf runs and their limited distribution within the detector, it was not possible to bin the runs in volume elements and, therefore, the runs have not been volume weighted. This means that the more extreme points will have a higher weighting than is appropriate given their position in the detector. In particular, an unrepresentatively large proportion of ^{252}Cf runs were near the bottom of the detector, so this region will be overweighted. Second, the R^3 distribution for neutron events drops off near the edge of the fiducial volume due to increased neutron capture on the acrylic vessel. This means that any effects near the edge of the volume, such as this scaling effect, will have a smaller impact on the neutron flux than they otherwise might have. For these reasons, the electron scaling was also taken to represent the neutrons for the FTP fitter.

5.7 Energy Dependence of the Fiducial Volume

The uncertainty on the fiducial volume is an important measurement because it can affect the flux of neutrinos reconstructing inside the fiducial volume. Even more important than this, however, is the energy dependence of this measurement, as this could affect the flux in an energy dependent manner and, hence, distort the extracted neutrino spectrum. The ^{16}N source could not be used for this measurement since, as discussed in chapter 2, it is a monoenergetic γ source, resulting in only a narrow range of electron energies. The ^{8}Li source was used instead, which samples a similar energy distribution to that of CC events. As before, any systematic differences observed between the data and Monte Carlo were parameterised and propagated to the signal extraction as systematic uncertainties.

5.7.1 Measurement Using ^{8}Li

Both the vertex shift and the reconstruction resolution can display biases with energy that would affect the neutrino flux inside the fiducial volume, so both must be taken into account in the measurement of the energy dependence. The parameter chosen, which would incorporate both of these effects simultaneously, was the fraction of events that

reconstructed inside the true source position:

$$\Gamma = \frac{\textit{Number of events reconstructing inside the source position}}{\textit{Total number of events}} \tag{5.6}$$

Considering multiple point sources on the boundary of the fiducial volume, it is evident that the value of Γ can be used as a measure of the number of events mis-reconstructing inside the volume. The effects of additional sources inside and outside the boundary should cancel, providing that they have the same value of Γ and extend to sufficient radii. The acrylic vessel is at a radius of $600\,\text{cm}$, so there is a $50\,\text{cm}$ buffer of detection medium outside the fiducial volume and the resolution of the detector is roughly $20\,\text{cm}$, so this assumption should be valid for neutrino events.

Once again, the important feature of this analysis was a comparison between data and Monte Carlo events to determine whether the effects were accurately reproduced. Γ was expected to have a value of roughly 0.5 with some variations due to solid angle effects, but these should be reproduced in the Monte Carlo simulation and, so, the only differences observed should be systematic energy-dependent effects. A first order polynomial was used to parameterise the energy dependence of the data to Monte Carlo ratio, as follows:

$$\frac{\Gamma_{data}}{\Gamma_{MC}} = \xi * T_{eff} + \epsilon \tag{5.7}$$

where T_{eff} is the reconstructed kinetic energy of an event. The gradient of this line, ξ, gives the required energy dependence. The value of the intercept is irrelevant to this study since it represents a constant offset between data and Monte Carlo, which was taken into account in the previous ^{16}N studies of the fiducial volume.

The ^{8}Li source was deployed in very few positions within the SNO detector and the statistics are fairly low for each run. In particular, no ^{8}Li runs were taken on the boundary of the fiducial volume, so the first thing to be investigated was any position dependence of the energy dependence, ξ, to determine whether runs at other locations within the detector could be reliably used for the measurement.

^{8}Li runs were taken at four radial positions within the detector. Small uncertainties in the source position would result in a global offset, not a differential effect with energy. The uncertainty on the stated source position can be on the order of $10\,\text{cm}$, as discussed previously. Therefore, the true source position was calculated for each run, independently for the data and the Monte Carlo, as the mean of the reconstructed position distribution. The same cuts discussed in previous sections were applied in this analysis to both the data and the simulation. Events were then separated into $1\,\text{MeV}$ bins according to their reconstructed energy. For each run, the number of events that passed all cuts and reconstructed inside the mean of the position distribution within each energy bin was counted and compared to the total number of events passing cuts in that bin, giving the

parameters Γ_{data} and Γ_{MC} as a function of energy for each run. The energy dependence of the fiducial volume was determined by finding the gradient of the best-fit line to a plot of the ratio of these parameters as a function of energy, giving a value for ξ in equation 5.7 for each run. Figure 5.13 shows the extracted energy dependencies for each individual run as a function of the radial position within the detector at which that run was taken.
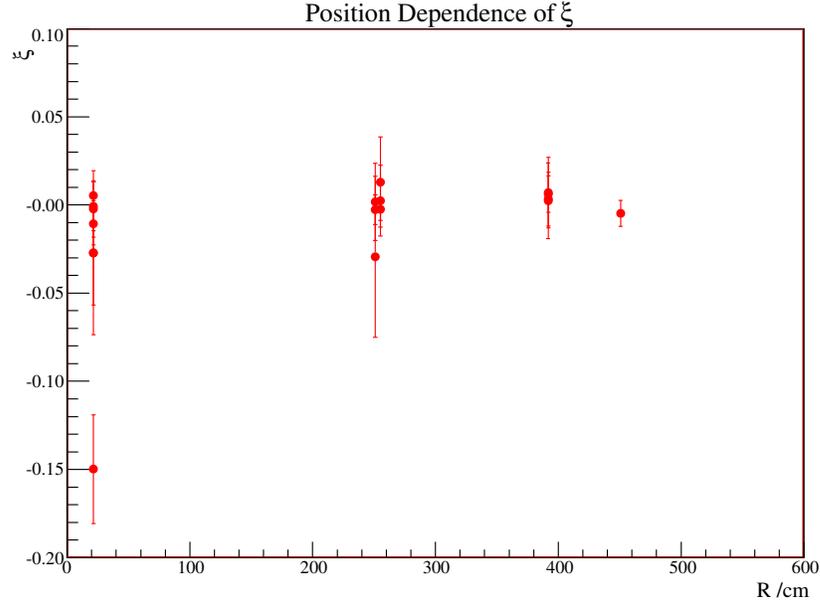
Figure 5.13: The energy dependence of the fiducial volume measurement (ξ from equation 5.7) as a function of position in the detector, using the FTP fitter in the salt phase.

This figure shows that, above a radius of about 200 cm, the energy dependence varies very little with position in the detector. More variation is observed for runs nearer the centre of the detector, which is to be expected since solid angle effects will skew the radial distribution and the assumption that the radius is much greater than the reconstruction resolution also breaks down. These runs are therefore more susceptible to source effects.

The uncertainties on each run were fairly large due to the limited statistics, so the lack of a position dependence seen in figure 5.13 was used to support the idea of combining all runs at radii greater than 200 cm to improve statistics for the analysis. The points within each energy bin were combined independently for both data and simulated events. The final results for Γ_{data}, Γ_{MC} and the ratio as a function of energy are shown in figures 5.14 and 5.15. A linear fit has been performed to figure 5.15, from 3.5-13.5 MeV in effective electron kinetic energy, to give the final measure of the observed difference of the energy dependence of the fiducial volume between the data and the simulation.

The results shown above were all for the FTP fitter in the salt phase; the full
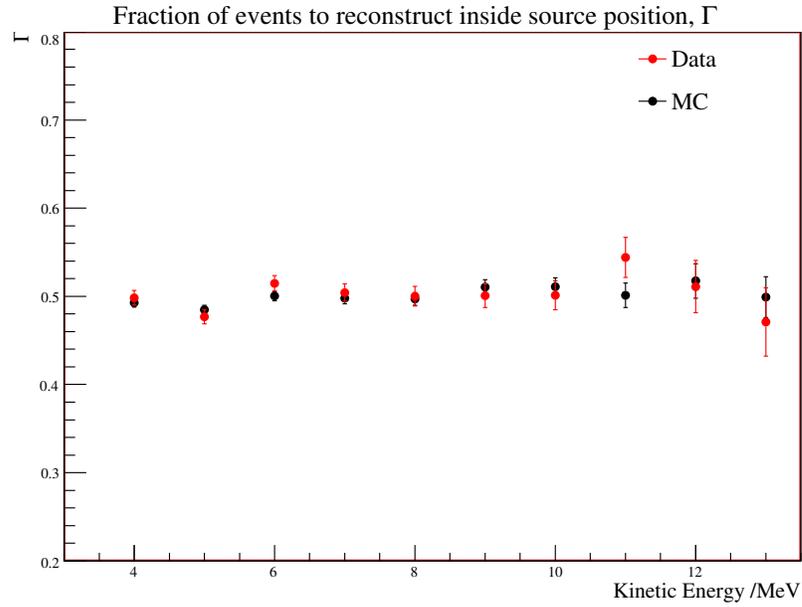
Figure 5.14: Energy dependence of fiducial volume measurement for data and Monte Carlo events, using the FTP fitter in the salt phase.
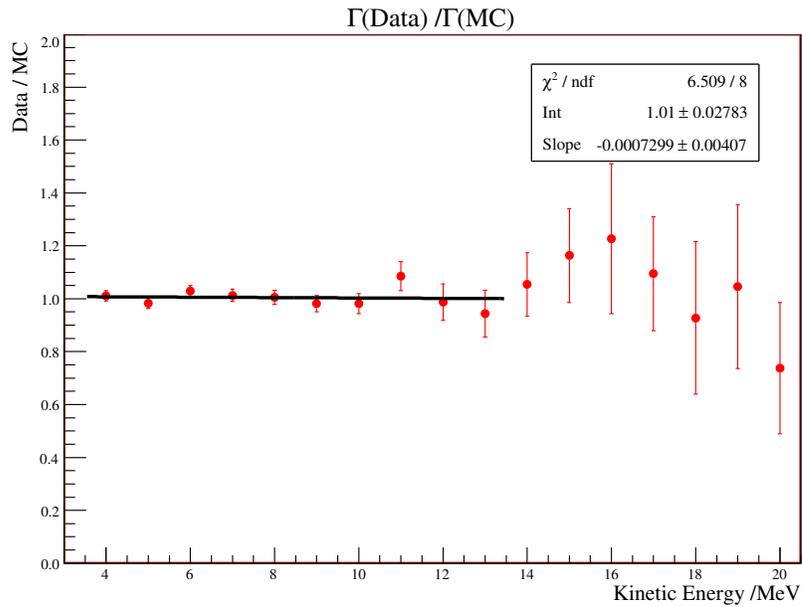
Figure 5.15: The difference between data and Monte Carlo results for the energy dependence of fiducial volume measurement, using the FTP fitter in the salt phase.

analysis was performed for both fitters in both phases and the final results obtained are summarised in table 5.5.

Energy Dependence of	D_2O		Salt	
Fiducial Volume /%	FTU	FTP	FTU	FTP
Energy dependence	-0.53 ± 0.44	0.36 ± 0.49	0.06 ± 0.40	-0.07 ± 0.41

Table 5.5: Difference between the energy dependence of the fiducial volume measurements for data and Monte Carlo events, for the FTU and FTP fitters in the D_2O and salt phases. Values are quoted as %.

5.7.2 Check Using Neutrino Data

The fiducial volume cut defined in section 3.3.3 limited the LETA analysis window to events reconstructing inside a radius of 550 cm. Therefore, the data outside this region could be used to provide an independent measurement of the systematic uncertainties. In particular, the data in a region centred around the position of the acrylic vessel, 560–640 cm, was used to verify the evaluation of the energy dependence of the fiducial volume described in section 5.7. In previous analyses an empirical function was fit to the radial profile of the neutrino events to fit for the position of the acrylic vessel [71]. The function used was a Hill function, of the form:

$$F(\rho) = A_0 + \frac{A_1 - A_0}{1 + \left(\frac{R_{\frac{1}{2}}}{\rho}\right)^{\beta}} \tag{5.8}$$

where ρ is the normalised, volume-weighted radial parameter introduced in chapter 3. The function is governed by four parameters: A_0 and A_1 designate the low and high radius limits, β determines the slope of the transition from low to high radius values and $R_{\frac{1}{2}}$ is the "half-point", which is taken to represent the radial position of the AV.

The full set of cuts were applied to both the data and the simulation to select a clean sample of neutrino events. Statistics in the data are very limited, so just three energy bins were used: 5.5-7.5, 7.5-9.5 and 9.5-19.5 MeV. A Hill function was fit to the events in each energy bin that passed all the applied cuts. Figure 5.16 is an example of one of these fits to the data.

To improve statistics in the Monte Carlo and to better model the data used, the three signal Monte Carlos were combined. The fluxes from the latest SNO publication [12] were used to combine the Monte Carlos in the ratio that would most closely represent the data. The ratio of the fitted position for the AV in the data and Monte Carlo was calculated in each energy bin and the result is shown in figure 5.17. This result measures a similar quantity to the ^{8}Li method, so the gradient of this plot can be compared to the energy dependence found in section 5.7.1. The neutrino data gives an energy dependence of 0.8 ± 2.1 % for the FTP fitter in the salt phase, which is in agreement with the value

Hill Fit to Data in 5.5-7.5MeV Bin

χ^2 / ndf	1.979 / 8
0	57.76 ± 3.39
1	131.7 ± 4.7
2	-407.3 ± 4057.6
HP	0.9633 ± 0.0328
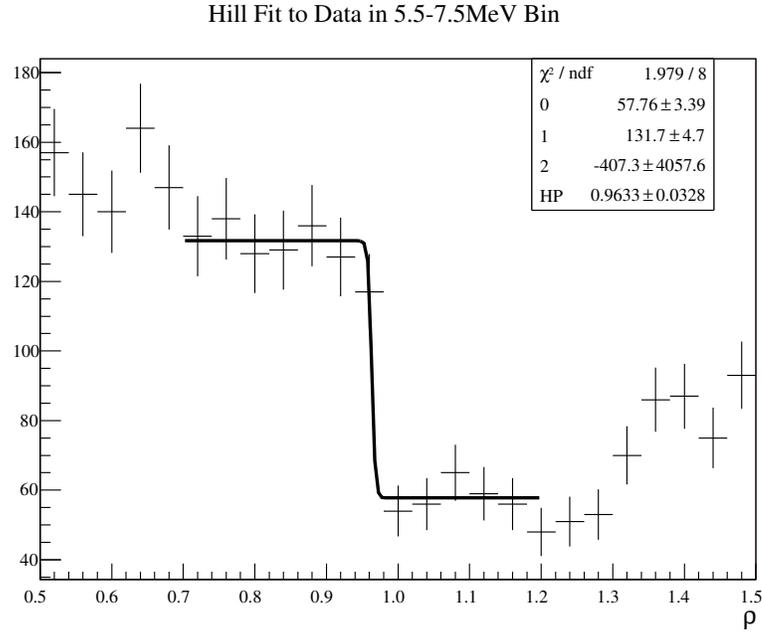
Figure 5.16: Example fit of a Hill function to the neutrino data in a 5.5-7.5 MeV kinetic energy bin, using the FTP fitter in the salt phase.

of -0.07 ± 0.41 % found from ^{8}Li , since both values are consistent with zero within a one sigma uncertainty. This is good supporting evidence that the use of the ^{8}Li source provided a reliable measurement of this effect.

5.8   Summary

A knowledge of the accuracy of vertex reconstruction in SNO is a vital ingredient to the evaluation of the neutrino fluxes in the final signal extraction. By lowering the energy threshold of the analysis, a large number of background events were included in the analysis window (described in section 3.4.3) many of which originated outside the central D_2O volume. In particular, events originating from the PMTs formed the largest single event type in the LETA data set. The reconstructed positions of these external background events follow a steep exponential distribution inside the fiducial volume (figures 3.10 and 3.11) and, thus, the reconstructed position is a powerful tool for separating out these events from neutrino events. As a result, the accuracy with which the position is known has a direct impact on the extracted fluxes.

This chapter has described the analysis of the systematic uncertainties associated with the process of reconstructing event positions within the detector. A measurement of the fiducial volume was of particular importance since this defines the number of targets

Data/MC Ratio for the Fitted AV Position
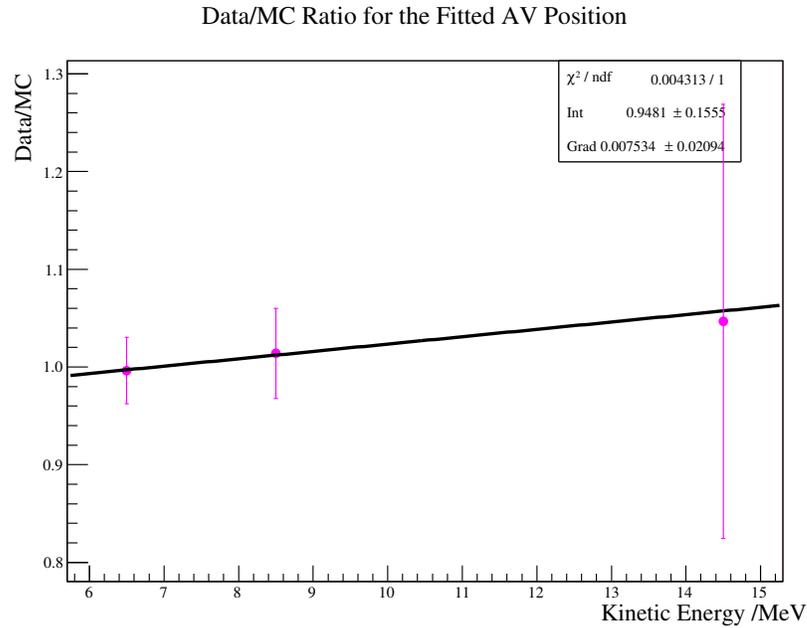
Figure 5.17: Energy dependence of the effective position of the AV, using the FTP fitter in the salt phase.

available for neutrino detection and, thus, has a direct impact on the flux measurement. A bias in reconstruction causing this number to differ from reality in an energy dependent fashion would have a direct impact on the energy dependence of the flux, distorting the extracted neutrino spectrum. This work has extended previous studies to the lower energy threshold. New methods have been developed, tested and implemented in order to deal with the new challenges presented. As a result, the knowledge of the detector has been improved and the uncertainties have been significantly reduced in comparison to previous values, thus increasing the accuracy with which the final neutrino fluxes and spectrum can be evaluated.

Chapter 6

Sacrifice

Removing background events from the raw data set in SNO involves applying a set of rejection criteria, known as data cleaning cuts. Events pass or fail these cuts based on their individual characteristics. Perfect signal-background separation is not typically achievable in any parameter and, so, some fraction of neutrino events also fail these cuts. The acceptance of signal events by the cuts is an important correction to the final neutrino fluxes; in particular, any energy dependence can affect the extracted spectral shape. The acceptance is calculated by measuring the sacrifice of signal events for a set of cuts, where the sacrifice is defined as the fraction of signal events failing the cuts. The acceptance is then given by:

$$Acceptance = 1 - sacrifice \tag{6.1}$$

This chapter describes the measurement of the sacrifice on the flux of solar neutrinos in the SNO data set, including the calculation of appropriate corrections and uncertainties to account for it.

6.1   Background Rejection

The LETA data set contains two main types of background events: physics backgrounds due to radioactive decays within the detector and instrumental backgrounds, which are caused by the detector itself. The combined rate of background events is significantly higher than the signal rate. Above 5 MeV, the instrumental background rate swamps other events and at lower energies the physics backgrounds dominate. Before the neutrino signal can be analysed, as many of these background events as possible must be removed from the data set. Two sets of cuts were developed to do this, as discussed in chapter 3. However, each set of cuts has an associated level of signal sacrifice, which must be taken into account in the final neutrino flux measurement.

The instrumental cuts involve a number of simple selections designed to remove

detector-related background events. These were applied once to the data, and the result was stored in a Data Analysis Mask Number (DAMN) bank [54]. A DAMN mask can be applied to any set of data in a bitwise manner to select only events that pass specific cuts.

The second set of cuts are known as the high level cuts, which were designed to accept signal events based on specific reconstructed characteristics of the events, such as energy, isotropy and timing. This suite of cuts was developed and extended specifically for the LETA analysis to better remove physics backgrounds [72]. In particular they target external radioactive decay events originating in the light water or on the PSUP whose origins have been misreconstructed inside the fiducial volume.

Each high level cut requires a reconstructed vertex in order to calculate the cut parameters. The reconstruction algorithms, or 'fitters', used to calculate this vertex position are discussed in more detail in chapter 5. An additional contribution to sacrifice comes from events that fail the reconstruction process due to fitter failure.

For the purposes of the LETA analysis, the possible sources of sacrifice were divided into three groups: fitter reconstruction, the instrumental cuts and the high level cuts.

6.2 Measuring the Sacrifice

The final extracted neutrino fluxes must be corrected for the measured level of sacrifice from the data cleaning cuts. As described in chapter 3, the Monte Carlo simulation was used to generate the PDFs used in the signal extraction procedure. These PDFs were used to predict the number and distribution of each type of event expected in the analysis window. To account for the level of sacrifice in the data set, a correction factor was applied to the Monte Carlo-predicted fluxes. For cuts that were not applied to Monte Carlo events and were only applied to data, this correction was given by the acceptance of the cuts in the data, reducing the Monte Carlo flux by the number of events that would have been removed had the cuts been applied. For cuts that were applied to the Monte Carlo distributions as well as to data, the correction was given by the ratio of the acceptance for data and Monte Carlo events:

$$\mathcal{R} = \frac{Data\ acceptance}{Monte\ Carlo\ acceptance} \tag{6.2}$$

Since a direct measurement of sacrifice cannot be made from the neutrino data set (if there was a way to tag individual neutrino events, then the majority of the analyses discussed in this thesis would be superfluous!) calibration source data and the associated Monte Carlo were used to provide answers. For many of the cuts, calibration sources were used to directly calculate the ratio of the acceptances. However, the mechanisms of many

of the instrumental backgrounds are not well enough understood for them to be modelled in the Monte Carlo simulation and, so, for these cuts a direct measurement of the sacrifice was made using the calibration data. The acceptance of the cuts was then calculated and applied as a correction to the Monte Carlo PDFs during the signal extraction procedure described in chapter 8.

The reconstruction algorithms and the high level cuts were applied to both data and Monte Carlo events and, so, some level of sacrifice was observed in both sets. In this case, the correction applied to the Monte Carlo flux was the ratio as defined above, $\mathcal{R}$.

Each source of sacrifice must be measured for each of the two fitters used for the LETA analysis, in both phases of SNO, for both electron and neutron-like events. In the D_2O phase, ^{16}N and ^8Li were used to represent both electron and neutron-like events. In the salt phase, ^{252}Cf was used for the neutrons. Although the ^{252}Cf source was deployed in the D_2O phase, the long capture time for neutrons makes the events very difficult to identify. Since neutron capture on deuterium results in a single γ of roughly $6.25\,\mathrm{MeV}$, this process is well modelled by the ^{16}N source. See section 2.3.2 for further description of the calibration sources.

The energy range considered for this analysis was that employed for the LETA analysis as described in chapter 3: 3.5–$11.5\,\mathrm{MeV}$ in effective electron kinetic energy. Virtually no physics backgrounds are expected above this value and the number of neutrino events also drops steeply. Corrections and uncertainties for energies above this range were conservatively assumed to be equal to the values in the highest energy bin (11–$11.5\,\mathrm{MeV}$). When the *'full energy range'* is referred to below, that refers to 3.5–$11.5\,\mathrm{MeV}$.

6.2.1 Event Selection

The sacrifice is a measure of the loss of *signal* events to the cuts, not backgrounds, so to make an accurate measurement of the sacrifice, a sample of pure signal events is required; background contamination in the event set could artificially inflate the measurement. In Monte Carlo studies this was clear-cut, since events were generated by event type, so the event sample was pure from the start. In data, more caution was required. For the ^{16}N and ^8Li sources, this was straightforward. Both these sources are tagged (see section 2.3.2) so genuine source events could be identified on an event-by-event basis. The ^{252}Cf source is not tagged, but as discussed in section 2.3.2, an algorithm has been developed to identify neutron events from the source. For the fitter sacrifice, all tagged events were considered as good events. For all further studies, a good event was required to also have a successfully reconstructed vertex and to reconstruct within the defined fiducial volume for the analysis (within a radius of 550cm) and inside the specific energy range discussed above.

6.2.2   Correlations

Correlations can exist between different cuts, so splitting the cuts into smaller sets for study leads to the possibility of double counting the sacrifice. Previous studies [54, 73] have shown that the instrumental and high level cuts have either very small correlations or are completely uncorrelated within the signal region. Therefore, these two sets of cuts can be treated independently. The instrumental cuts are completely uncorrelated with reconstruction. The high level cuts require the successful reconstruction of a vertex, but these can be treated as independent if the high level cut sacrifice is considered as being the sacrifice *after* a successful vertex has been reconstructed.

6.2.3   Differential Sacrifice

As well as an absolute measure of the sacrifice in the data set and in the Monte Carlo PDFs, the differential sacrifice must be considered. Since the LETA analysis involves extracting the neutrino flux using the event distributions in terms of both energy and position, the sacrifice as a function of these variables must be understood to reduce possible sources of systematic error. In particular, the sacrifice as a function of energy is a vital measurement, since this will have a direct impact on the extraction of a neutrino spectrum from the data set. If the number of events accepted by the cuts differs in data and Monte Carlo in an energy-dependent fashion, then this would affect the extracted signal flux differently at different energies, distorting the extracted spectrum. For this reason, the energy and radial dependence of all sources of sacrifice were explored in depth. The radial dependence was studied in terms of the normalised radial parameter, ρ, introduced in chapter 3, which naturally weights the events according to detector volume.

Another feature that must be studied is the stability of the sacrifice measurements since data from SNO extends over a period of several years. Any variation with time could introduce a bias into the final results, so a selection of source runs spanning the full period of data taking were used to investigate this possibility.

For the instrumental cuts, these features were studied in the sacrifice of calibration data. Since the fitters and the high level cuts were applied to Monte Carlo events as well as data, any differences between these features for the two event sets must be understood. For example, a time dependence in the sacrifice of data events that was accurately modelled in the Monte Carlo distribution would not affect the final flux extraction, but a position dependence that was not reproduced would need to be taken into account.

6.3 Reconstruction Sacrifice

Reconstruction is the process of estimating the position at which an event took place within the detector. The 'fitter' algorithm uses individual PMT information to perform this calculation. An event fails reconstruction if the fitter fails to return a valid vertex for the event. Events can fail reconstruction in one of two ways: either the fitter can fail to compute a vertex or the vertex it returns can be at an unphysical fit position (such as out in the mine rather than inside the detector). Therefore, the two constraints placed to define a successful fit were the successful return of a vertex and a fit position within 12m of the centre of the detector. FTU rarely fails to return an event vertex, but it can reconstruct positions outside this 12m limit. As discussed in section 3.3.2, FTP automatically fails an event if the seed vertex is outside the PSUP and, accordingly, fails to return a vertex more often, but a higher proportion of the successfully reconstructed vertices are within the detector.

The fitters were run on both the full data set and the Monte Carlo events so, as discussed above, calibration sources could be used to find the ratio of the acceptance in the data and the Monte Carlo, $\mathcal{R}$, and the associated uncertainties.

A primary measurement of fitter sacrifice was generally made using the ^{16}N source. The ^{8}Li source was used to provide an energy dependent check and ^{252}Cf was used to measure the sacrifice for neutrons in the salt phase.

The nature of the fitter sacrifice means that the events that fail (*i.e.* the sacrificed events) do not have a successfully reconstructed vertex and, therefore, have no value for reconstructed position or energy. Differential sacrifice for this measurement was therefore evaluated in terms of the generated event position (or the stated source position for data events) and the number of hit phototubes for the event, known as the *'NHit'* value. Any difference between the source and the reconstructed positions should be taken into account in the reconstruction uncertainties (see chapter 5).

6.3.1 Sacrifice and Uncertainties

Figure 6.1 shows the fitter sacrifice as a function of NHit for the FTU fitter in the D_2O phase for ^{16}N data and Monte Carlo events. The resulting acceptance ratio, $\mathcal{R}$, is also shown, along with the best fit first order polynomial. The ^{8}Li source was used to study the energy dependence, so the gradient of this fit was fixed to zero to just return a weighted average of the points. For ^{16}N data, the ratio for both fitters in both phases fit to exactly 1.0 to better than a 0.01% significance. Therefore, the ratio was taken to be 1.0, *i.e.* no correction was applied for the fitter sacrifice. However, uncertainties on this number still need to be taken into account. These are described below, including investigations into

both the energy and position dependence of the sacrifice.
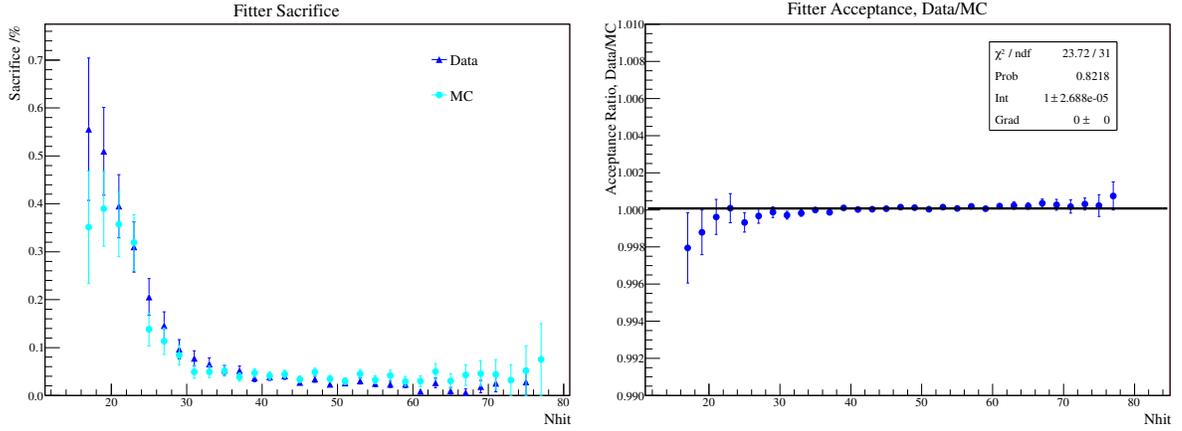
Figure 6.1: Fitter sacrifice in data and Monte Carlo (left) and the resulting acceptance ratio (right) as a function of NHit for the FTU fitter in the D_2O phase, using the ^{16}N source.

Higher energy events generate more light, resulting in a greater number of hit phototubes and, so, more information available to reconstruct event vertices. The ^{16}N source gives a measure of the sacrifice at low energies and it was therefore expected to give an upper bound on the global sacrifice value. The ^8Li source was used to give an idea of the energy dependence of the fitter sacrifice. Figure 6.2 shows the absolute sacrifice and the acceptance ratio for the ^8Li source using the FTU fitter in the D_2O phase. A comparison of figures 6.1 and 6.2 clearly illustrates the lower sacrifice for the ^8Li source, although uncertainties on the ^8Li measurement are greater due to lower statistics. The best fit line shown in figure 6.2 illustrates the NHit dependence of the acceptance ratio. When allowed to vary in the fit, the gradient of this line fit to 0.001%. This dependence on NHit was taken to be insignificant and the gradient was therefore fixed to zero for the fit, so the result is essentially a weighted average of the points. The ratio for the ^8Li source did not always fit out to exactly 1.0. Any discrepancy from this value was taken as a systematic uncertainty and the returned fit uncertainty was taken as a measure of the statistical uncertainty. The same method was used to find uncertainties for neutrons in the salt phase using the ^{252}Cf source.

It is also possible for the sacrifice to vary with position in the detector. The ^8Li and ^{252}Cf sources were deployed in a limited number of positions whereas the ^{16}N was more widely deployed throughout the detector volume, so ^{16}N was used for this measurement. The detector was split into equal volume radial bins and the acceptance ratio was found for the runs falling in each bin. Figure 6.3 shows the result for the FTU fitter in the salt
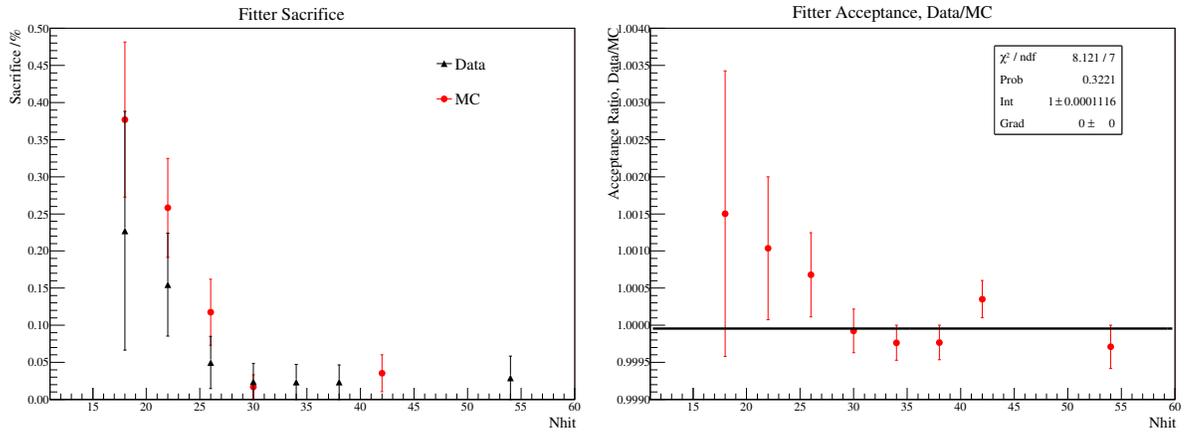
Figure 6.2: Fitter sacrifice in data and Monte Carlo (left) and the resulting acceptance ratio (right) as a function of NHit for the FTU fitter in the D_2O phase, using the ^8Li source.

phase. The median deviation of these points from the best fit value was used as a measure of the position-related uncertainty.
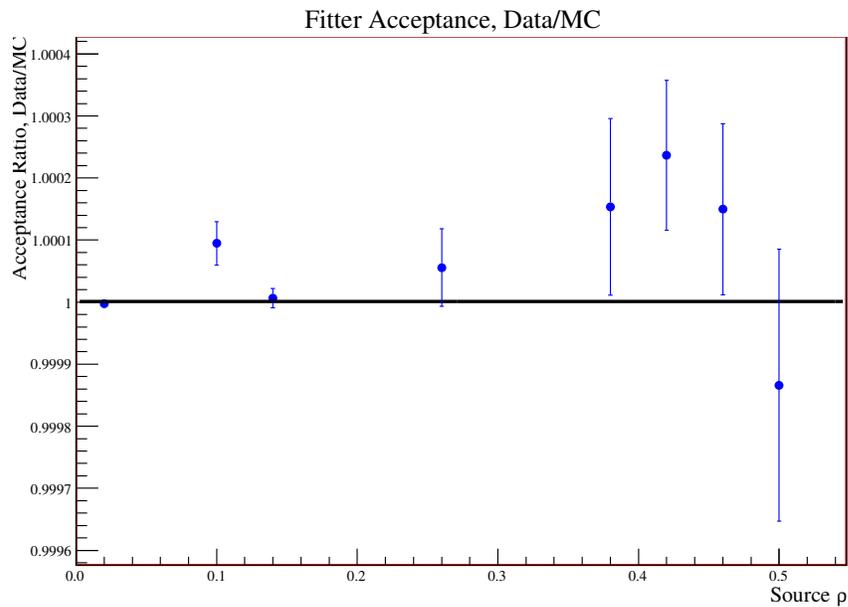
Figure 6.3: Ratio of the acceptance in data and Monte Carlo as a function of ρ for the FTU fitter in the salt phase, using the ^{16}N source.

6.3.2 Results

No significant differences were seen in the application of the fitters to data versus Monte Carlo events and, so, no correction was deemed necessary to account for differences in fitter sacrifice. The sacrifice of the fitters for each of the three signal types (CC, ES and NC) was determined for each fitter in each phase using signal Monte Carlo events. Results are given in table 6.1.

The uncertainties evaluated in the previous section were added in quadrature to return the final uncertainty on the acceptance ratio; values for each source of uncertainty are quoted in table 6.1. The statistical uncertainty (stat uncert) is the fit uncertainty from the energy-dependent fit and the energy dependence refers to the deviation of that fit from 1.0. ^{8}Li was used for electrons and for neutrons in the D_2O phase and ^{252}Cf was used for salt neutrons. The position dependence was derived from the ^{16}N fit and was applied to both electron and neutron events. This uncertainty was combined with the uncertainties on the signal sacrifice to give the final uncertainty due to fitter sacrifice.

Parameter	D_2O		Salt	
/%	FTP	FTU	FTP	FTU
CC sacrifice	0.007 ± 0.000	0.136 ± 0.002	0.007 ± 0.000	0.152 ± 0.002
ES sacrifice	0.034 ± 0.002	0.237 ± 0.004	0.031 ± 0.001	0.244 ± 0.004
NC sacrifice	0.229 ± 0.006	0.536 ± 0.009	0.019 ± 0.001	0.117 ± 0.002
Stat uncert (e)	0.034	0.011	0.037	0.023
Stat uncert (n)	0.034	0.011	0.000	0.008
Energy-dep (e)	0.060	0.000	0.090	0.040
Energy-dep (n)	0.060	0.000	0.000	0.100
Position-dep	0.010	0.002	0.009	0.008

Table 6.1: Reconstruction sacrifice in the SNO data set, with statistical (stat) and systematic uncertainties, for electron (e) and neutron (n) events, in %.

6.4 Instrumental sacrifice

Instrumental, or "non-physics" background events in SNO are a serious problem, as their event rate is much higher than that of neutrino events. However, these instrumental backgrounds have very different topologies, charge and/or time distributions compared with physics events, which allows them to be removed from the data. The DAMN mask is a collection of cuts that remove instrumental backgrounds within SNO (see [54] for a description of the individual cuts). Some of these cuts fail events based on the individual characteristics of that event, such as the charge on individual PMTs, while others examine

a larger data set and remove a period of time that fails certain criteria. An example of the latter is a retrigger event, which is a second trigger of the detector caused by a single event. After a triggered event there is a deadtime in which further triggers cannot occur. A 1–2 ms delayed pulse from a large event can cause a second detector event. After-pulsing inside the PMTs can have the same effect: when photoelectrons are created at the photocathode they create positive ions, which drift backwards and can knock further electrons off the photocathode, causing a second signal delayed by the ion drift time. This after-pulsing has a lower charge than the original signal, but for a large event it can be enough to trigger the detector. For the measurements of sacrifice described here, all such time-based cuts were removed from the DAMN mask. The sacrifice of these cuts is affected by individual run conditions; they were tracked run-by-run and the effect was taken into account in a livetime correction (see section 8.2.2). Certain other cuts were removed from the DAMN mask for the baseline measurements, for specific reasons, and these are discussed in sections 6.4.2 and 6.4.3 below.

The DAMN mask was applied to data only, not to the simulated events and, so, the correction for instrumental sacrifice comes from a direct measurement of the sacrifice in the calibration data. The ^{16}N , ^{8}Li and ^{252}Cf sources were used for this measurement.

6.4.1 Baseline Sacrifice Measurement

As in previous studies [74], some problems were encountered with the ^{16}N source when trying to use it to measure the sacrifice in the data. The measurement consistently returned higher values than the other sources, as illustrated in figure 6.4, and also revealed a noticeable feature in the energy dependence of the sacrifice, shown in figure 6.5. Checking each cut individually revealed that this feature was due to the AMB (analogue measurement board) cut, as was seen previously. The analogue measurement board measures the total charge deposited in an event independently of channel properties, thus avoiding problems with badly calibrated channels. The cut is very sensitive to event timing, as it measures the charge deposited in a specific time window, determined to be roughly the time it takes for light to traverse the detector. The cause of the observed energy-dependent oscillation in the sacrifice may be linked to this sensitivity to event timing, or it may be due to the bin width intrinsic to the ADC. This appeared to be a problem localised to the ^{16}N source, but despite plausible theories, the explicit reason for it was never discovered. For this reason, the ^{8}Li source was used to measure the instrumental sacrifice along with its energy dependence for electron-type signals and ^{252}Cf was used for neutrons. Since the difference between the two electron-type sources was never satisfactorily explained, this difference was included as an uncertainty in the final sacrifice calculation.

An example of the energy-dependence of the sacrifice of the instrumental cuts mea-
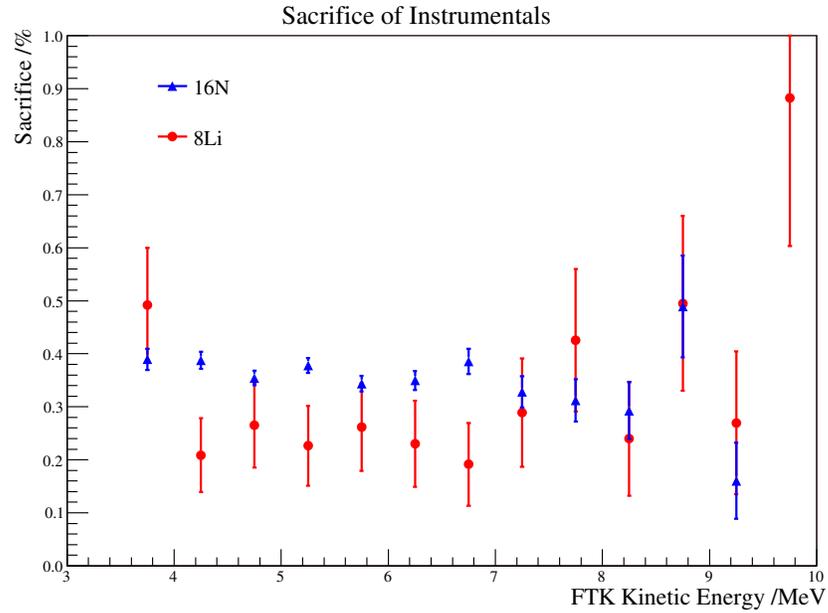
Figure 6.4: Instrumental sacrifice as a function of energy for the FTK/FTP fitters in the D_2O phase, comparing the ^{16}N and ^8Li sources.
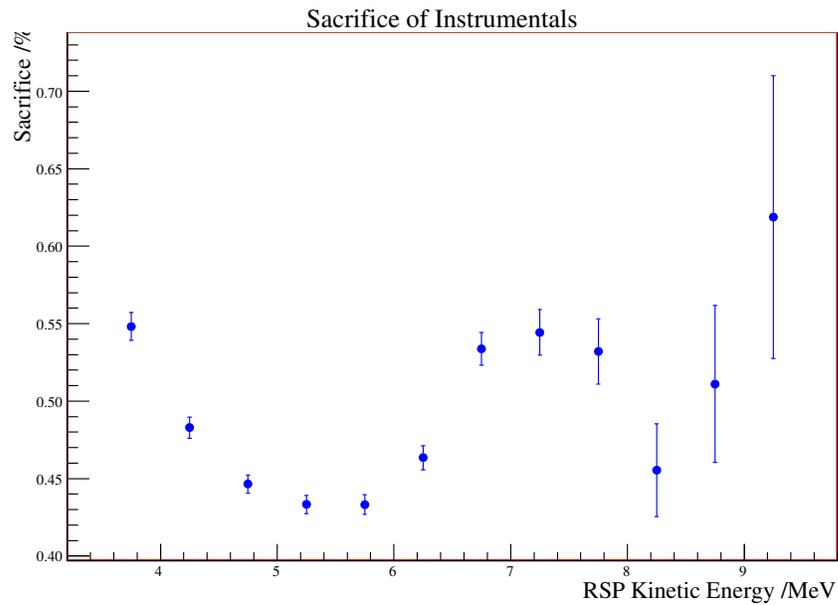
Figure 6.5: Instrumental sacrifice in ^{16}N data as a function of energy for the RSP/FTU fitters in the salt phase.

sured using the ^8Li source is shown in figure 6.6. The uncertainties shown in each bin are purely statistical. This illustrates the stability of the sacrifice across the energy range so,

instead of applying a correction in each bin individually, the weighted mean of the sacrifice in all bins from 3.5–11.5 MeV was calculated and applied as a global correction with the associated statistical uncertainty. A systematic uncertainty was included to account for integrating over the full energy range. The median deviation of the points in each energy bin from the weighted mean was used as a measure of this uncertainty.
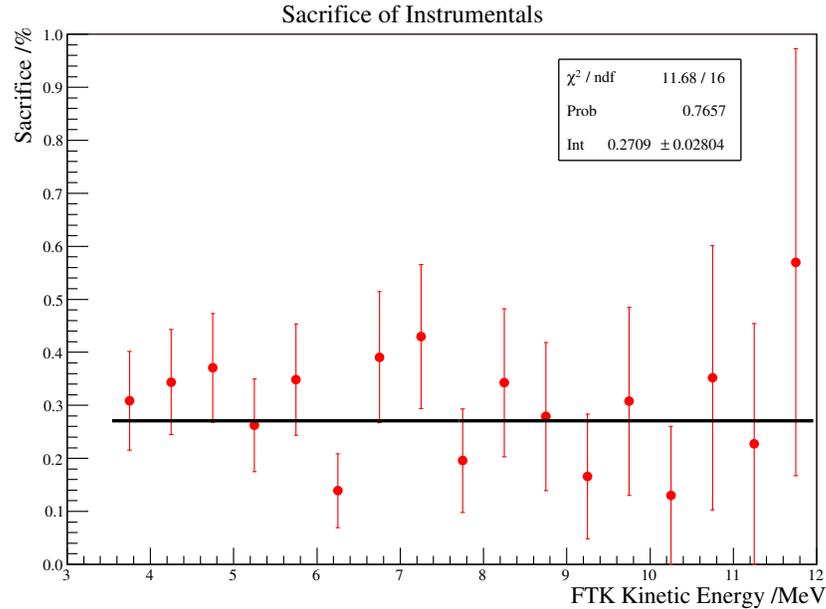
Figure 6.6: Instrumental sacrifice in ^{8}Li data as a function of energy for the FTK/FTP fitters in the salt phase. The best fit value is shown, along with the fit statistics.

The global instrumental sacrifice calculated using each source, for each phase of SNO and using each of the two fitters is shown in table 6.2. Although the instrumental sacrifice should not be directly affected by the choice of position fitter, each position fitter was teamed with a different energy fitter for the purposes of the LETA analysis, which can affect the sacrifice via the energy limits imposed during event selection. Several corrections to these numbers were necessary and these are discussed in the following sections. In addition, further systematic uncertainties due to the position and time dependence of the sacrifice were investigated and are presented below.

The ^{16}N source was used to measure the position dependence due to a wider sampling of source positions compared with ^{8}Li. One ^{16}N scan was selected for each phase, taken over three days in the D$_2$O phase and two in salt, to study the position dependence of the sacrifice. This was to avoid double counting other uncertainties in the sacrifice, such as time variations, by integrating across the entire data set. The detector was divided into equal volume radial bins. Events were allocated to these bins according to their

	Sacrifice /%			
Calibration	D$_2$O		Salt	
Source	FTP	FTU	FTP	FTU
^{16}N	0.362 ± 0.006	0.363 ± 0.006	0.466 ± 0.003	0.467 ± 0.003
^{8}Li	0.272 ± 0.026	0.273 ± 0.025	0.270 ± 0.028	0.295 ± 0.029
^{252}Cf	—	—	0.282 ± 0.003	0.281 ± 0.003

Table 6.2: Instrumental sacrifice in the SNO data set, with statistical uncertainties.

reconstructed position and the sacrifice of events in each radial bin was then calculated. As previously discussed, a known offset had been observed between the ^{16}N and ^{8}Li source results for instrumental sacrifice. This has already been taken into account by a systematic source uncertainty so, to avoid double counting, the measure of the position-related uncertainty in the sacrifice was taken to be the median deviation of the position-dependent ^{16}N results from the best fit value *to those points*, not to the global ^{8}Li best-fit. An example of the deviation of the sacrifice with position is shown in figure 6.7.
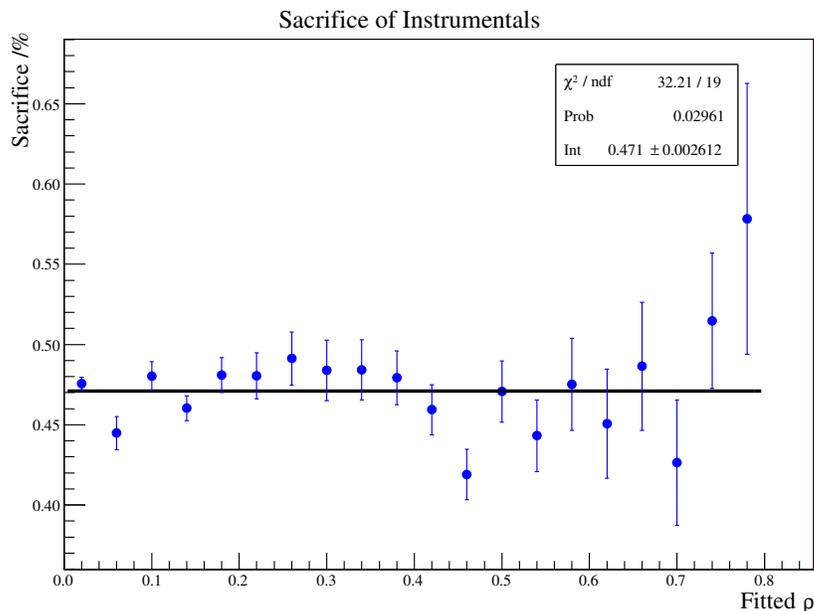
Figure 6.7: Instrumental sacrifice in ^{16}N data as a function of ρ for the FTK/FTP fitters in the salt phase. The best-fit global value for the sacrifice is indicated by the horizontal black line.

In a similar manner, the stability of the sacrifice measurement was determined by binning ^{16}N runs according to their Julian Date and finding the median deviation from the best-fit value, as shown in figure 6.8. For this measurement, a selection of central runs

was chosen spanning the full period of data-taking in each phase. No strong trends with time were observed. The deviation from the central value was larger than that observed for the position dependence, as demonstrated by the poor χ^2 value, but this was taken into account in the associated uncertainty.
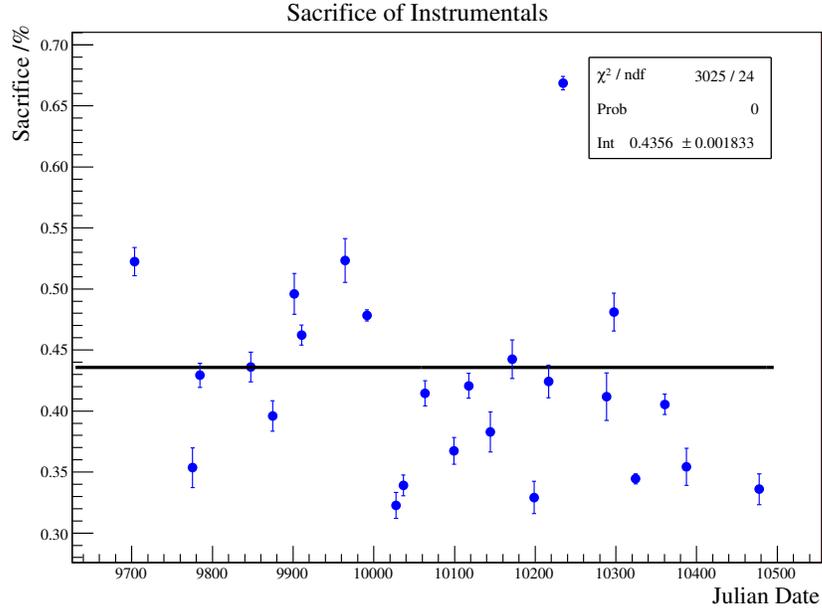
Figure 6.8: Instrumental sacrifice in ^{16}N data as a function of Julian Date for the FTK/FTP fitters in the salt phase. The best-fit global value for the sacrifice is indicated by the horizontal black line.

The values of all the systematic uncertainties derived above are summarised in table 6.3. The values derived from ^{8}Li measurements were used to represent the uncertainty on the sacrifice for all electron-type signals (CC and ES events) and for neutron signals in the D$_2$O phase. ^{252}Cf measurements were used for neutron signals in the salt phase. For the position and time dependence of the sacrifice, the ^{16}N measurements were used to represent all signals.

6.4.2 The OWL Sacrifice

The OWL cut vetos events based on the number of hits seen on the outward-looking tubes, thus removing events that generated significant levels of light outside the detector [54]. This cut was removed from the DAMN mask for the baseline sacrifice measurements discussed in section 6.4.1 since the sacrifice can be measured directly from data.

The OWL sacrifice was measured using the pulsed global trigger (pulseGT), which was fired at a rate of 5Hz throughout data taking to record the state of the detector.

Systematic	D_2O		Salt	
Uncertainty /%	FTP	FTU	FTP	FTU
^{8}Li source uncertainty	0.090	0.090	0.197	0.172
^{252}Cf source uncertainty	—	—	0.184	0.186
^{8}Li energy dependence	0.009	0.037	0.039	0.030
^{252}Cf energy dependence	—	—	0.008	0.004
^{16}N position dependence	0.009	0.006	0.007	0.0003
^{16}N stability	0.010	0.009	0.013	0.014

Table 6.3: Systematics uncertainties on instrumental sacrifice in the SNO data set. All values are given as absolute uncertainties in the global sacrifice result (where the sacrifice is measured in %). All uncertainties quoted here are applied as two-sided symmetric errors on final sacrifice results and therefore only the magnitude of the uncertainty is given.

The resulting information was useful for diagnostic purposes and noise rate studies. The sacrifice of the OWL cut is dominated by coincidences between multiple noise hits on the outward-looking tubes. Events only fail the OWL cut due to random accumulation of sufficient hits in the OWL system, so the sacrifice can be measured by finding the proportion of events in the pulseGT data set that fail the cut.

The OWL sacrifice was calculated separately for each phase, with the uncertainty given by Poisson statistics. The results are given in table 6.4 below. This additional sacrifice was added to the sacrifice results given in table 6.2, with the uncertainties being combined in quadrature.

Phase	OWL Sacrifice /%
D_2O	0.0334 ± 0.0002
Salt	0.0207 ± 0.0001

Table 6.4: OWL sacrifice in the SNO data set, with statistical uncertainties.

6.4.3 Correction for Bad Channels

In the original DAMN mask there were two cuts known as the 'charge' and 'cluster' cuts, respectively. The charge cut was designed to fail any event with a very high or railed charge and the cluster cut failed events in which there were hits on four or more adjacent tubes in electronics space, implying some level of crosstalk between the channels. The latter cut contained an inherent bias due to the correlation between electronics space and physical space: an event near the edge of the vessel would be more likely to fail the cut.

During the D_2O phase, one of the electronics boards in the detector developed a

fault in the ADC that caused the PMT charge to saturate for each hit on 32 channels. The charge cut particularly suffered from this fault, with the sacrifice measuring roughly 20% in runs containing the bad board. To increase the robustness of this cut and to improve the bias in the cluster cut, the two were combined. The resulting QCluster cut was designed to look for occurrences of very high charge on a single PMT within a cluster of adjacent hit channels. This change significantly reduced the sacrifice but there was still a noticeable systematic effect from the bad board.

A second instrumental cut affected by the railed charges was the QvT cut. The QvT cut was designed to cut events by looking at the PMT with the highest charge. An event with a particularly high charge on that tube relative to the median across all hit tubes and an early hit time compared to the other tubes would fail the cut.

The effect of the bad board was an artificial inflation of the sacrifice, specifically for the QVT and QCluster cuts. This was corrected by removing all the hits on the bad card in the software and measuring the systematic effect of the change ([54], [75]), which led to a reduction in the baseline of the sacrifice measurement. The bad board was not present for all runs and, so, a further correction was made to take into account the resulting time dependence. The PulseGT was used to monitor the effect and the integrated time variation was added back in to the sacrifice measurement. The resulting corrections and associated uncertainties are summarised in table 6.5. These corrections were not applied in the salt phase, since the ADCs were fixed towards the end of the D_2O phase.

Since the stability of the QvT and QCluster cuts in the D_2O phase has been taken into account here, these cuts were not included in the DAMN mask for the general D_2O stability studies discussed in section 6.4.1.

An additional systematic uncertainty on the baseline instrumental sacrifice measurement of -0.021% was introduced due to electronics effects, which were flagged by the ECA (electronic calibration). The saturated channels didn't flag the ECA since it flagged badly calibrated rather than broken channels, so there is no overlap with the previous corrections. This is a one-sided uncertainty that needs to be included with the other systematics. Full details of the calculation are described in [54].

Effect	Sacrifice Correction /%
Inflation of QvT	-0.088 ± 0.007
Inflation of QCluster	-0.03 ± 0.004
Stability of QvT & QCluster corrections	0.027 ± 0.002
ECA uncertainty	-0.021

Table 6.5: Corrections to the instrumental sacrifice in the D_2O phase due to the bad board, taken from [54], and other electronics effects.

6.4.4   Results

The correction due to instrumental sacrifice was calculated by combining the sacrifice values from table 6.2 with the OWL sacrifice from table 6.4 and the additional bad board corrections in table 6.5. The values obtained from the ^{8}Li source were used as the correction for electron events and for neutron events in the D$_2$O phase. The ^{252}Cf source was used for salt phase neutron events.

The uncertainty on this correction for each signal was obtained by adding all the individual uncertainties in quadrature, including the statistical uncertainties on the global sacrifice and on the OWL sacrifice, the uncertainties on the bad board corrections and the systematic uncertainties given in table 6.3.

6.5   High Level Cut Sacrifice

Previous to the LETA analysis, a small set of high level cuts existed that were used to discriminate between signal events and physics backgrounds arising from radioactive decay within SNO. Due to the lowering of the energy threshold for LETA, background events are now much more prevalent in the analysis window. This is due to the steep exponential rise in the energy spectrum of both bismuth and thallium decay events at low energies. Therefore, for the purposes of this analysis the high level cuts were developed and extended [72] to reduce external backgrounds. Attention was also given to reducing any energy bias in the cuts, which could lead to distortions of the extracted CC or ES spectra.

It was found that different cuts were more effective on events reconstructed using different fitting algorithms and therefore two sets of cuts were developed: one for the FTK/FTP analysis and one for the RSP/FTU analysis. The sacrifice of each set of cuts is considered here.

The event features exploited by the high level cuts to reject backgrounds are well modelled in the Monte Carlo and, so, these cuts were applied to both data and simulated events. However, the nature of sacrifice is such that it is a measure of the tail of a distribution. The accuracy with which the Monte Carlo reproduces data in the tail is not guaranteed to be equivalent to that in the main body of the distribution. For this reason, calibration source measurements were used as a normalisation for the sacrifice measurement. The ratio of the acceptance of the high level cuts in data and Monte Carlo was found using the calibration data and was applied as a correction to the signal Monte Carlo events.

As before, the ^{16}N, ^{8}Li and ^{252}Cf sources were used for the majority of the sacrifice measurements. However, some caveats exist for the high level cuts. ^{16}N is a γ source,

so the Čerenkov light that is detected is generated at a finite distance from the source, at the point at which the generated γ Compton scatters an electron. ^{8}Li is an electron source, which means the events occur very close to the source itself. This means that the source blocks almost all the backward-going light, some of which is then reflected back out, altering the hit distribution of the event. Several of the high level cuts are sensitive to this distribution, in particular the cut on event isotropy. Therefore, extra care must be taken if the ^{8}Li source is to be used for these measurements.

6.5.1 An Iterative Procedure

In previous studies, the high level cut sacrifice was on the order of 1% [54]. A small increase on this number was expected due to the extra developments made to the cuts in order to reject low energy background events. The lowering of the analysis energy threshold could also have introduced biases into the spectrum of the cuts that were previously outside the analysis window. It was particularly important to minimise biases in the spectrum to avoid distortions in the final extracted CC and ES spectra.

A first look at the sacrifice of the new suite of high level cuts highlighted some problems. Figure 6.9 shows the sacrifice of the cuts for Monte Carlo ES electron events. As is apparent from this plot, a strong bias was observed in the new cuts at both low and high energies.
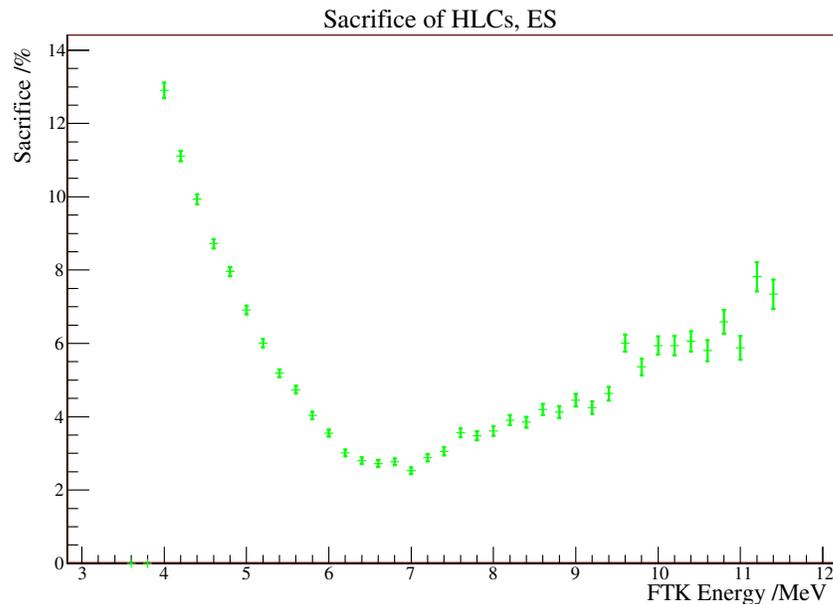
Figure 6.9: Sacrifice of the high level cuts for ES events in the D$_2$O phase, using the FTK/FTP fitters.

Each cut was investigated individually to determine which were causing the observed spectral biases. Two in particular were identified as being in need of readjustment. This is illustrated in figure 6.12, which also demonstrates the improvements resulting from the subsequent tuning [72], discussed below. This problem was observed to only be significant in the FTK/FTP analysis; no readjustment was required for the RSP/FTU analysis.

ITR Adjustment

The first cut in need of readjustment was the In-Time Ratio (ITR) cut. The ITR value for an event is the ratio of the number of prompt hits to the total number of hits. A prompt hit is defined as one with a time residual of between -2.5 and 5.0 ns, where the time residual is the difference in the PMT hit time from the predicted time of flight from the reconstructed vertex to the PMT. Negative values are possible due to time jitter in the PMT response and the finite reconstruction resolution of the vertex fitter algorithms.

The ITR cut was one of the original set of high level cuts used in past SNO analyses. Some degree of bias at low energies was seen in the sacrifice, but it was negligible above the chosen energy threshold and, therefore, unimportant. For the first iteration of the LETA high level cuts, the cut value was kept unchanged and the same low energy bias was observed. This was to be expected since low energy events have fewer PMT hits and, therefore, less information available to reconstruct the event position. The spread of reconstructed positions is, therefore, broader, which has a direct impact on event timings and leads to a greater spread in the ITR distribution. This is illustrated in figure 6.10.

The original format of the cut was to reject all events with an ITR value less than 0.55. The observed increase in the width of the ITR distribution at low energies caused more signal events to reconstruct outside this range and was, therefore, the direct cause of the observed spectral bias in sacrifice. A significant bias in a cut is more of a problem if the sacrifice is high. For example, a 50% bias in a cut with a sacrifice of only 0.01% would be negligible. The aim of the adjustment of ITR was therefore to reduce the overall sacrifice as well as the bias in the spectrum.

ITR is a ratio of hit numbers and, so, the distribution was expected to be binomial. As the NHit of an event increases, more information is available to determine the mean of the distribution, so the width of the distribution narrows. Conversely, at low energies the spread in the distribution is broader. As shown in figure 6.10, the mean of the distribution is roughly constant across the energy range ($\mu \sim 0.74$). The width of the distribution is given by:

$$\sigma = \frac{\sqrt{NHit * \mu * (1 - \mu)}}{NHit} \sim \frac{0.43}{\sqrt{NHit}} \tag{6.3}$$

A new cut was devised to cut at a fixed number of σ below the mean ITR value [47],
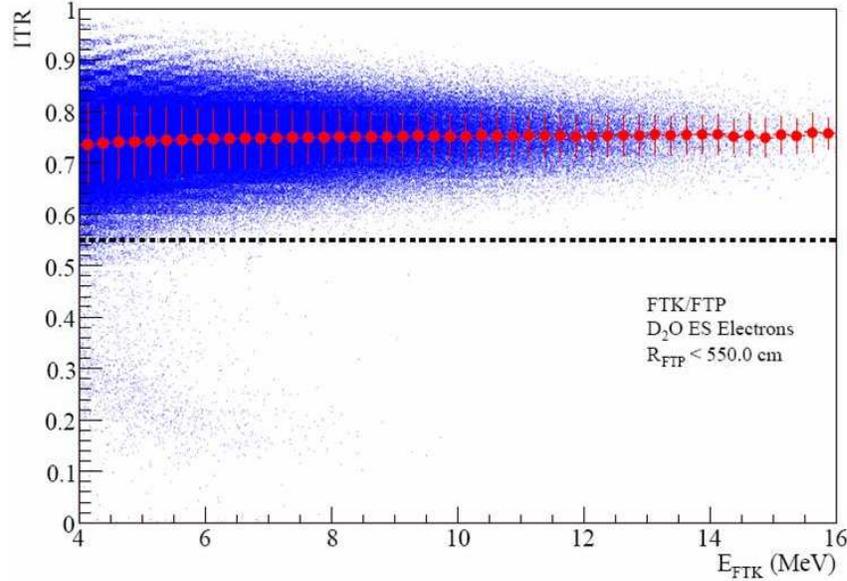
Figure 6.10: ITR distribution as a function of reconstructed energy for ES electrons in the D_2O phase, using the FTK/FTP fitters. The red points and associated error bars show the mean and width of the ITR distribution. Figure courtesy of Joshua R. Klein.

using the NHit dependence of σ given in equation 6.3. This loosened the cut at low energies, such that fewer signal events were mistakenly rejected and the overall sacrifice was subsequently reduced, along with the spectral bias.

FTK Uncertainty Cut Adjustment

The FTK energy fitter uses a maximum likelihood technique (similar to that described in chapter 7) to find the most likely value of an event's energy given the number of hit PMTs. Each event is assigned asymmetric uncertainties in the reconstructed event energy based on the goodness of fit during the minimisation process. The high level cuts exploit these uncertainty values to reject poorly reconstructed events.

The main culprit of the biases at high energy were these FTK uncertainty cuts. However, there are several caveats where these cuts are concerned. The position fitters do not take into account the source geometry when reconstructing event vertices, so it is possible for FTP to reconstruct an event close to the source and pointing directly at it. Such an event is physically improbable. FTK takes the source geometry into account, so it assigns these events a greatly overestimated value of energy to compensate for the tubes it believes were shadowed by the source and the assigned uncertainty values are similarly large. Therefore, these events would be cut by the FTK uncertainty cuts and appear to contribute to sacrifice. However, this is a source effect, not a true representation

of the topology or distribution of signal neutrino events and, so, the effect should not be incorporated into a measurement of the neutrino sacrifice. To remove these events before calculating the sacrifice, any event that reconstructed within 50cm of the centre of the source and with a momentum vector pointing inwards towards the source (an angle between the reconstructed event direction and the directional vector from the source to the event of less than 90°) was discounted.

Even with this extra event selection criterion, a significant high energy bias was still present, as evident in figure 6.12. A readjustment of these cuts was therefore required. The original format of the cuts used an energy-dependent energy resolution, $\sigma(E)$, to reject all events that had an uncertainty value greater than a fixed multiple of σ for either the upper or lower energy uncertainties. The parameterisation of the Čerenkov photon distribution in FTK uses different parameter values for discrete energy bins across the full SNO spectrum of $2 - 60\,\mathrm{MeV}$ [51]. One of the boundaries occurs near the middle of the LETA energy range, around $7\,\mathrm{MeV}$. For this reason, the first iteration of the uncertainty cuts [76] was to introduce a first order energy dependence into the cut value and to calculate and apply it in two distinct regions as defined by this boundary. In addition, the uncertainty cuts were restricted to only act on events below $13\,\mathrm{MeV}$. To reduce the complexity inherent in this cut, it was replaced by a quadratic function of energy that was applicable across the full energy range, up to $13\,\mathrm{MeV}$ [47]. Both this final version and the original cut value are shown in figure 6.11. The effect of this retuning was to significantly reduce the high energy bias in the sacrifice, as illustrated in figure 6.12.

FTK Uncertainty Cuts - Persistent Issues

Despite the significant reduction in the spectral bias, problems were still evident in the uncertainty cut sacrifice. In particular, steeper energy trends in the ^{16}N source data and a much stronger dependence on reconstructed position within the detector were observed for these cuts in comparison to the rest of the high level cuts. For this reason, the decision was made to separate the FTK uncertainty cuts from the remainder of the high level cut set and to treat them independently. For this to be a valid action, the two sets of cuts must be shown to be uncorrelated.

Consider a scenario with two sets of cuts and a total of N events, of which n_1 fail the first set, n_2 fail the second set and n_{12} fail both sets. The expected statistical overlap of the sacrifice of the two sets of cuts would be:

$$\alpha = \frac{n_1}{N} \times \frac{n_2}{N} \tag{6.4}$$

Therefore the value of $\frac{n_{12}}{N}$ can be compared to α to determine whether there are any significant correlations between the sets. Table 6.6 shows the values obtained using ^{16}N
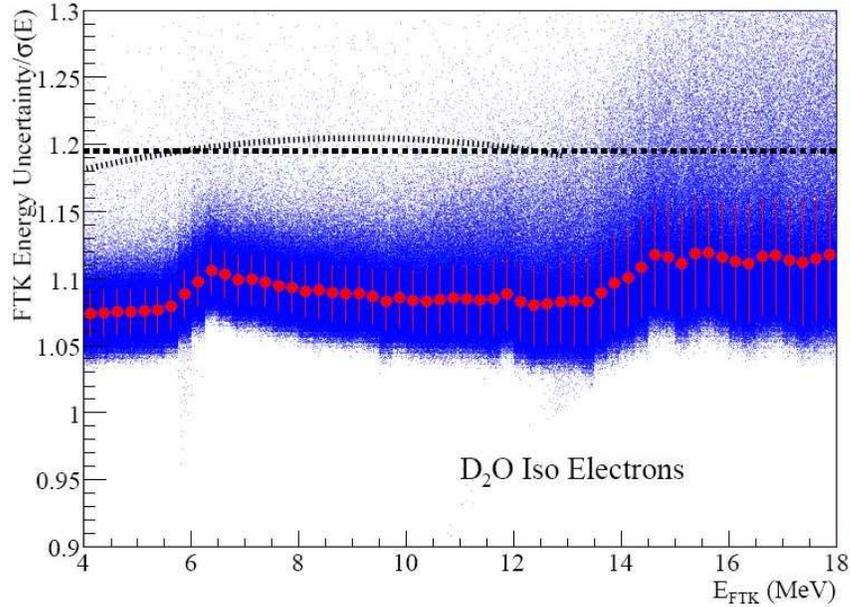
Figure 6.11: Distribution of the upper FTK energy uncertainty in terms of energy resolution ($\sigma(E)$) as a function of reconstructed energy for D_2O phase isotropic electrons using the FTK/FTP fitters. The horizontal dashed line and the curved dotted line represent the original and new cuts respectively. The red points and associated error bars show the mean and width of the distribution. Figure courtesy of Joshua R. Klein.

data and Monte Carlo. On the basis of this evidence, it was agreed that there was no significant correlation and the two sets could legitimately be treated as independent.

	Predicted overlap, α /%	Actual overlap, $\frac{n_{12}}{N}$ /%
D_2O phase data	0.039	0.050
D_2O phase Monte Carlo	0.037	0.022
Salt phase data	0.018	0.021
Salt phase Monte Carlo	0.014	0.014

Table 6.6: Correlations between the FTK uncertainty cuts and the remainder of the high level cuts using [16]N data.

For the remainder of this chapter, the term *'high level cuts'* now refers to the set of high level cuts with the FTK uncertainty cuts removed. Although the sacrifice in the Monte Carlo of both sets of cuts has been reduced to acceptable levels, the ratio of the acceptance for data and Monte Carlo ($\mathcal{R}$) still needs to be determined. The associated analysis for the truncated set of high level cuts is discussed in section 6.6 and the FTK uncertainty cuts in section 6.7.

Resulting Improvements

The effect of the adjustments of the ITR and FTK uncertainty cuts is clearly illustrated in figure 6.12. The sacrifice of the FTK/FTP set of high level cuts both before and after the retuning is shown, highlighting the original cause of the spectral biases and the resulting improvements. Figure 6.13 shows the new sacrifice of the high level cuts and the FTK uncertainty cuts for ES Monte Carlo events. Comparing this to figure 6.9, the improvement of the retuning is clear.
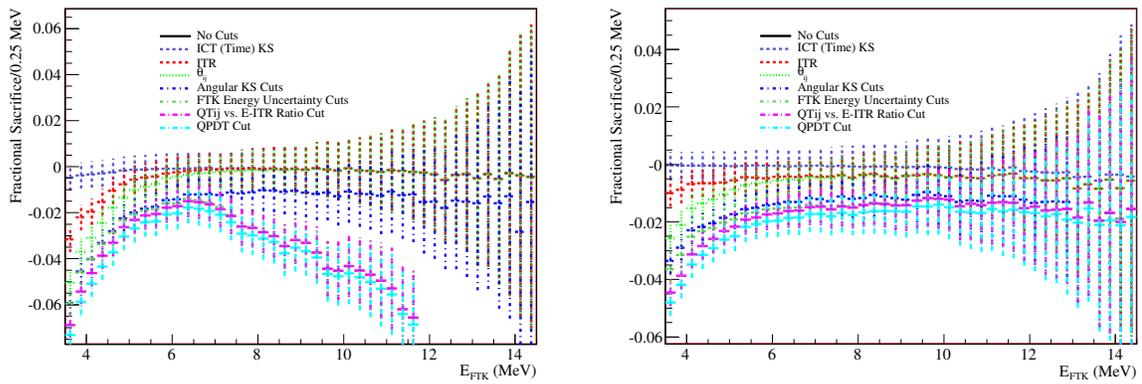
Figure 6.12: Cumulative sacrifice of the high level cuts for ES electrons in the D_2O phase, using the FTK/FTP fitters, before (left) and after (right) the high level cut retuning. Figures courtesy of Joshua R. Klein.
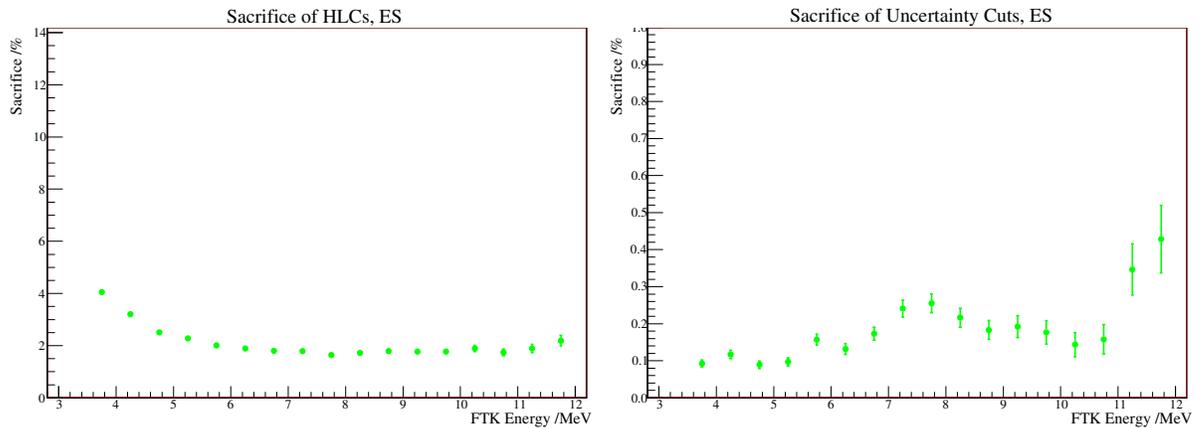
Figure 6.13: Sacrifice of the high level cuts (left) and the FTK uncertainty cuts (right) after retuning, for ES events in the D_2O phase, using the FTK/FTP fitters.

6.6 High Level Cut Acceptance Ratio

As discussed in section 6.5, the sacrifice of the high level cuts in the simulation must be corrected by any differences observed between the behaviour of the cuts on data and simulated events. The ratio of the acceptance of the high level cuts in the data and the simulation ($\mathcal{R}$) can be found using calibration source data and applied as a correction to the Monte Carlo simulation of signal neutrino events. In particular, any dependence of this ratio on observables such as position or particularly energy must be carefully studied to avoid introducing biases into the final neutrino flux measurements.

6.6.1 Calculation of Acceptance Ratio

A first iteration of the high level cut analysis showed that there was some radial dependence to the acceptance ratio. For this reason, the detector was split into seven bins using the normalised radial parameter, ρ. The first 6 bins were all of width 0.1 in ρ but, due to a deficit of runs near the edge of the fiducial volume, the final bin was given a width of 0.3 to increase statistics. This takes the analysis out beyond the edge of the fiducial volume purposefully to allow for reconstruction offsets and resolution.

The main measurement was made using the ^{16}N source. The acceptance ratio for each run was found in 0.5 MeV kinetic energy bins from 3.5–11.5 MeV. In order to appropriately combine the results from individual runs, each energy bin was considered separately. First, the error-weighted mean of all the runs in the same volume bin was found (using only the events falling into the energy bin under consideration) along with the associated statistical uncertainty. Then the results for each volume bin were combined in a volume-weighted fashion.

Figure 6.14 shows the acceptance ratio of the high level cuts as a function of energy for both the ^{16}N and ^{8}Li sources. For ^{16}N, the ratio is very flat at the lower energies but begins to drop off after about 9 MeV. This may well be due to the fact that ^{16}N is a monoenergetic source, so the higher energy events are often due to unusual sources such as pile-up of multiple events. This is a feature in the data that is not reproduced in the simulation, causing the data sacrifice to be higher and hence the acceptance lower than in the simulation and, therefore, causing the ratio to drop, as observed. If this is the genuine cause of the drop, then it is a source effect and not a feature that should be reproduced in the measurement of the sacrifice of neutrino events. The ^{8}Li source was used to confirm this theory. Although, as previously mentioned, care must be taken when using the ^{8}Li source for sacrifice measurements due to the source shadowing problems, a comparison of the ^{8}Li data to the results from the ^{16}N source below 9 MeV (in figure 6.14) shows good agreement between the two. Despite the poorer statistics, the drop that was seen

in the ^{16}N data at higher energies is clearly not present in ^{8}Li, suggesting that it was in fact a source effect unique to the ^{16}N. For this reason, instead of applying the ratio as a correction in each energy bin, a weighted mean was used over the first 10 energy bins of the ^{16}N data, from 3.5–8.5 MeV. This value was used for the correction at all energies.
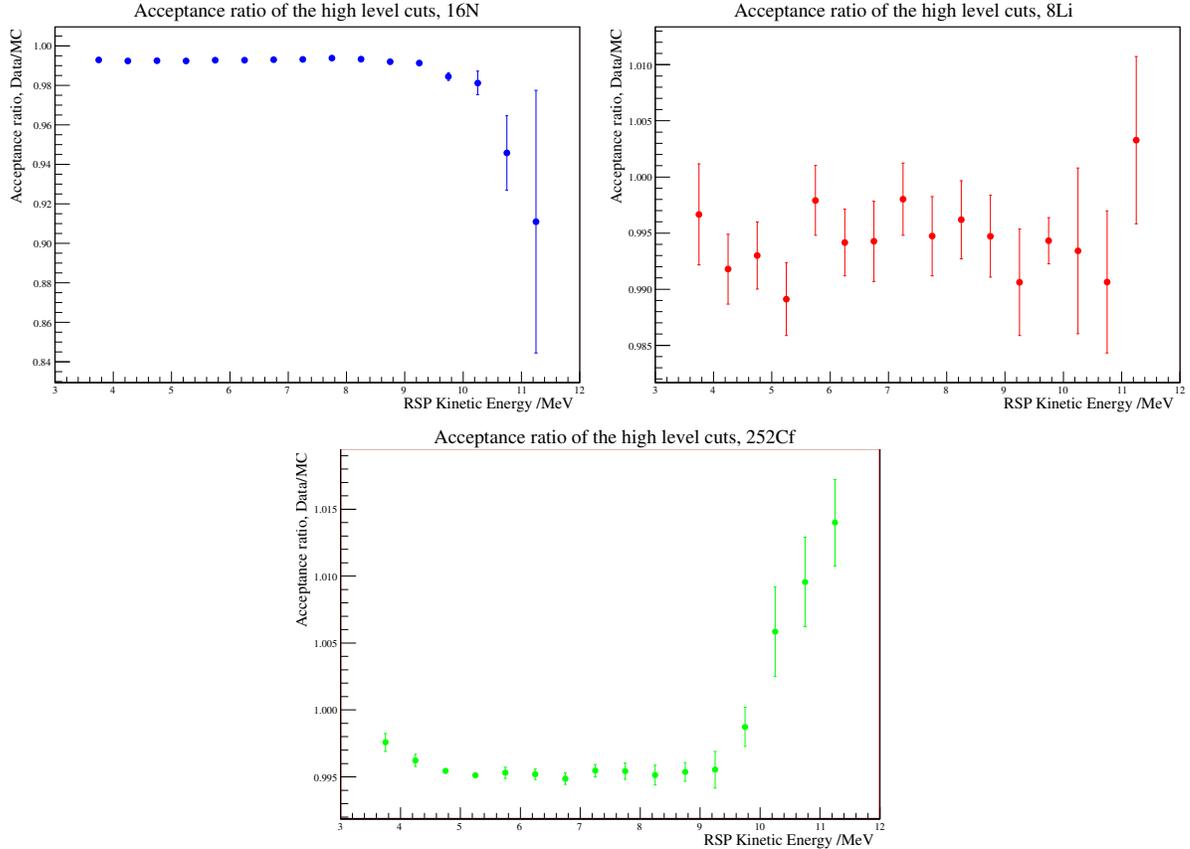
Figure 6.14: High level cut acceptance as a function of energy for the FTU fitter using the ^{16}N source (top left) and the ^{8}Li source (top right) in the D$_2$O phase and the ^{252}Cf source in the salt phase (bottom).

For the neutrons in the salt phase the case was slightly different, as illustrated in the third plot in figure 6.14. Again, the ratio was not flat across the full spectral range. In this case, the ratio actually rose slightly at both the very low and very high ends of the spectrum. One of the problems with the ^{252}Cf source is that source gammas can contaminate the data sample. These are not simulated in the Monte Carlo code and, so, they can cause extra events in the data set which can be accepted by the cuts and, therefore, increase the data acceptance relative to the simulation. These events are likely to be at low energies, but can also occur in coincidence with neutron events and so affect the high energy tail as well. If the shapes in this spectrum are due to source

gammas then, again, it is a source effect that should not be taken into account in the neutrino sacrifice measurement. Although the QBurst algorithm was applied to remove these source gammas (section 2.3.2), it is possible that some still exist in the data. To test this, radial cuts were placed at first 0.5 m and then 1 m from the centre of the source to remove those events occurring close to the source. The neutron capture distance in the salt phase is greater than the γ scattering length and, so, this is likely to remove most of the source gammas but a smaller fraction of the neutrons. The rise at lower energies was entirely removed by these cuts and the higher energy rise was also affected. This is supporting evidence that the features were due to source gammas. For this reason, a weighted mean was used over the energy range 4.5–9.5 MeV for the correction at all energies. The discrepancy from flat was taken into account in the uncertainties discussed below.

6.6.2   Uncertainties

Although the correction was calculated as a global value, the uncertainties on it are given in an energy dependent fashion, in 0.5 MeV bins in kinetic energy. This is because some variations in energy were averaged over or discounted when calculating the correction and their effect still needs to be taken into account.

The statistical uncertainty was given by the uncertainty on the specific fit used to determine the global correction value. For electron-like events this came from ^{16}N data and for neutrons it came from ^{252}Cf.

The assignment of an energy-dependent systematic uncertainty to the acceptance ratio was handled differently for electron and neutron-like events due to the different methods of calculating the correction:

Energy Dependence for Electron Events

The acceptance ratio in the ^{16}N dropped off at energies above 8 MeV, but the ^{8}Li data remained roughly flat, implying that the drop was due to event types specific to the ^{16}N source. Therefore, the uncertainty was determined by calculating the weighted mean of the ^{8}Li points across the full energy range and finding the median deviation from this value. This was applied as a two-sided uncertainty.

Energy Dependence for Neutron Events

For neutron events, only one calibration source exists: the ^{252}Cf. Some discrepancies were seen in the ratio at low and high energies, which were thought to be due to source gammas. However, the evidence for this was not absolutely conclusive and, since the

effect of different radial cuts varied the shape of the acceptance ratio with energy, this was incorporated into the uncertainty. In each bin, the difference between the global correction value and the value in that bin was taken as a one-sided uncertainty.

Stability Measurements

The ^{16}N source was used for determining the stability of the high level cut acceptance ratio across the full time span of data-taking due to the greater prevalence of available runs. The stability of the ratio was found by binning central ^{16}N runs according to their Julian Date and then finding the median deviation from the best-fit value, as illustrated in figure 6.15.
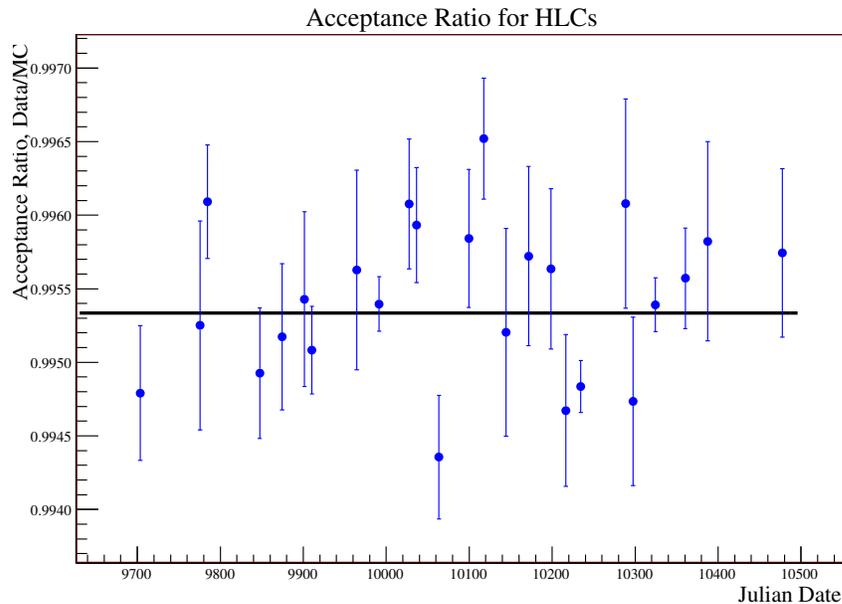
Figure 6.15: High level cut acceptance as a function of Julian Date for the FTK/FTP fitters in the salt phase, using the ^{16}N source. The black line represents the weighted mean.

Radial Uncertainties

As previously mentioned, a weak radial dependence was observed in the acceptance ratio for the high level cuts. The ^{16}N source was used for this measurement due to its wider deployment within the detector. Runs were divided up into volume elements, as defined in section 6.6.1. Figure 6.16 shows the dependence of the ratio on the run position.

To account for the variations, the runs within each volume element were combined (integrating over all energies between 3.5 and 11.5 MeV) and a volume weighted best-fit
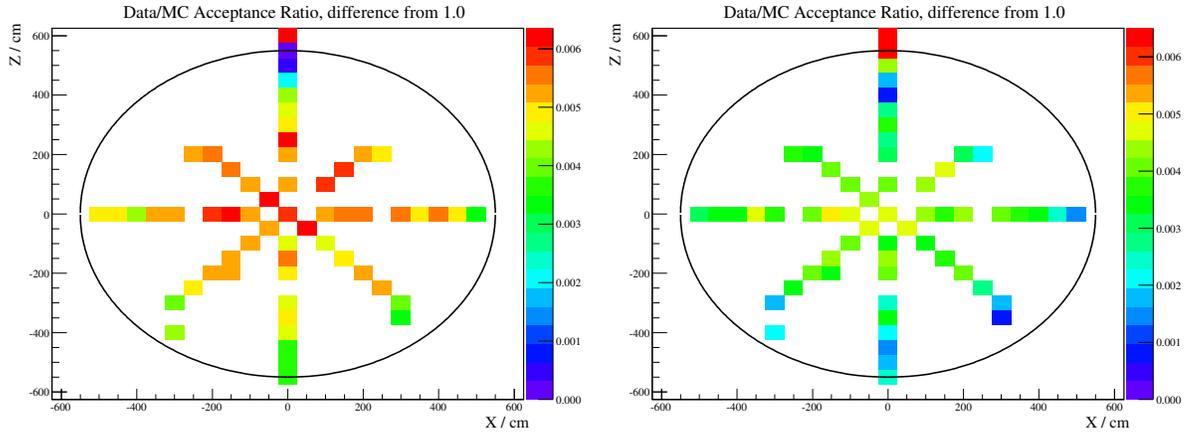
Figure 6.16: Discrepancy of the high level cut acceptance ratio from 1.0 as a function of position within the detector for the RSP/FTU (left) and FTK/FTP (right) fitters in the salt phase, using the ^{16}N source. The black line represents the boundary of the fiducial volume.

value was found. Previously, the uncertainty for measurements of this sort was estimated using the median deviation from the best-fit value. This provides a measure of the spread around the best-fit value without being pulled by any tails or extreme points. However, in this case the known shape of the distribution needs to be taken into account to accurately include the effect of the tail, so the volume weighted mean deviation from the best fit value was used. Figure 6.17 illustrates this. The ^{16}N source was used to calculate this systematic uncertainty for both electrons and neutrons since the ^{252}Cf source was only deployed in a limited number of positions and a comparison at those points suggested that the position dependence was similar for both sources.

6.6.3   Results

The values for the corrections to be applied to the simulation due to the ratio of the high level cut acceptance for data and Monte Carlo events are given in table 6.7. The electron correction was taken from the ^{16}N fit and was applied to electron events and neutron events in the D$_2$O phase. The neutron correction was taken from the ^{252}Cf fit and was applied to salt phase neutrons only.

The values of each of the uncertainties discussed above are also given in table 6.7, except the energy uncertainty for salt phase neutrons, which was calculated individually for each energy bin and can best be judged from figure 6.14. The stability and radial systematic uncertainties discussed above were applied constantly across the energy range. This equates to applying the same uncertainty value in each energy bin. The individual
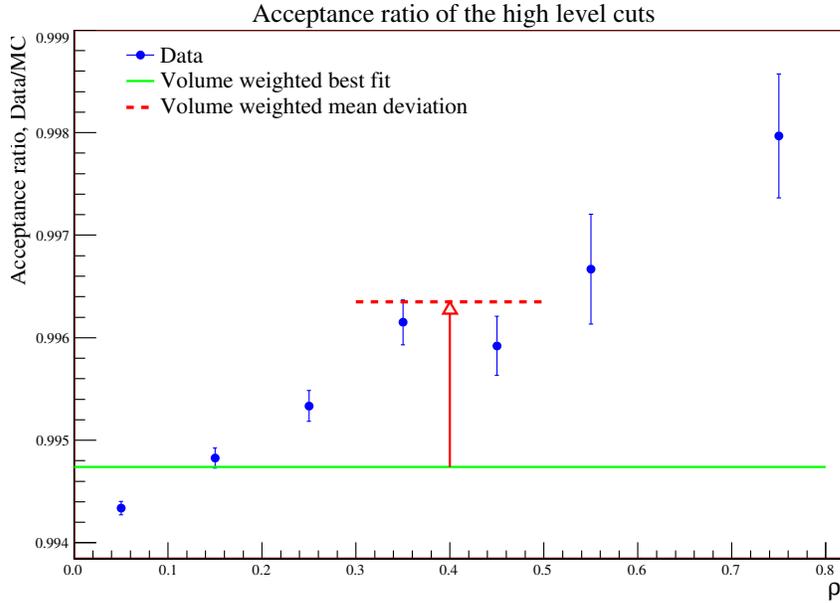
Figure 6.17: High level cut acceptance ratio as a function of position for the RSP/FTU fitters in the salt phase, using the ^{16}N source.

uncertainties were combined in quadrature in each energy bin to give the final uncertainty on the acceptance ratio. This uncertainty was then combined with the statistical uncertainty on the acceptance of the cuts in Monte Carlo on a signal-by-signal basis to give the final uncertainty on the correction.

	D$_2$O		Salt	
	FTP	FTU	FTP	FTU
Electron correction	0.9945	0.9928	0.9959	0.9945
Neutron correction	0.9945	0.9928	0.9983	0.9953
Stat uncert (e) /%	0.0273	0.0228	0.0159	0.0114
Stat uncert (n) /%	0.0273	0.0228	0.0196	0.0115
Energy-dep (e) /%	0.1897	0.0049	0.1226	0.0261
Energy-dep (n) /%	0.1897	0.0049	E-dep	E-dep
Position-dep /%	0.1630	0.0898	0.3144	0.1618
Stability /%	0.0805	0.0298	0.0130	0.0144

Table 6.7: High level cut sacrifice in the SNO data set, with statistical (stat) and systematic uncertainties, for electron (e) and neutron (n) events.

6.7 FTK Uncertainty Cut Acceptance Ratio

6.7.1 Calculation of Acceptance Ratio

As discussed in section 6.5, the FTK uncertainty cuts were separated out from the rest of the high level cuts for the FTK/FTP analysis. By their very nature, the sacrifice of the uncertainty cuts should be identical for data and Monte Carlo events under the assumption that the parent distributions are the same. FTK uses only the reconstructed position and direction of an event, along with the total number of hit tubes, to determine an event's energy and the associated uncertainties. There is no dependence on the hit distribution. For this reason, no correction was applied for the uncertainty cuts - the acceptance ratio was taken to be 1.0. However, there are still uncertainties associated with this, which were explored using similar methods to the high level cuts.

6.7.2 Uncertainties

By taking $\mathcal{R}$ to be 1.0 for the uncertainty cuts, an assumption has been made that the distributions are flat in all observables. Systematic trends in the ratio are discussed below. As a measure of the statistical uncertainty, the spread in the points from a flat distribution was calculated. The weighted mean of the value in each energy bin was taken and the error in this mean was used to represent the statistical uncertainty. ^{16}N data was used for electron events and ^{252}Cf data for neutrons, applying the same limited energy ranges that were used to calculate the global correction for the high level cuts in section 6.6.1. The reason for not using the full energy range was that both ^{16}N and ^{252}Cf showed some non-uniformity with energy at the extreme ends of the spectrum, which was taken into account in the energy related systematic uncertainty and should not be double counted.

The stability of the acceptance ratio across the data set was calculated, as before, by finding the median deviation of the acceptance ratio from the weighted mean for a set of central ^{16}N runs binned in Julian Date.

Neck Effects

The position dependence of the uncertainty cuts is more complicated than for the remaining high level cuts. The acceptance ratio for events near the neck varied much more significantly from the expected value of 1.0 than events elsewhere in the detector. Figure 6.18 shows the magnitude of the variation in the ratio from 1.0 with position in the detector. The spherical detector has essentially been flattened along the y-axis, showing only the x and z-axes, but it was confirmed that the y-axis matched the x-axis very closely. There is a very clear peak in the discrepancy of the acceptance ratio from 1.0 near the

neck of the acrylic vessel (this is, in fact, a drop in the value of the ratio due to inflated sacrifice in the data near the neck). The neck represents a clear break in the symmetry of the otherwise spherical vessel. Any mismodelling of the neck in the Monte Carlo simulation could easily cause the detector response to differ from the data, leading to the observed behaviour in the acceptance ratio.
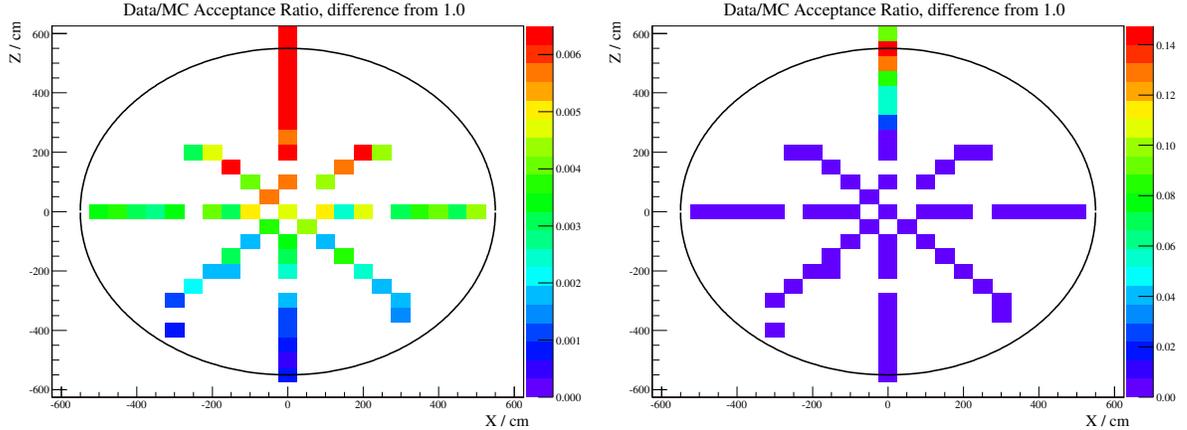
Figure 6.18: Magnitude of the difference of the FTK uncertainty cut acceptance ratio from 1.0 as a function of position in the detector in the salt phase, using the ^{16}N source. The scale in the left figure is normalised to show the variations in the body of the detector. The scale on the right shows the full extent of the discrepancy at the neck.

Several investigations were performed to investigate this effect. Five ^{16}N runs along the z-axis of the detector were selected to enable a comparison of data to the simulation at a range of positions. The distribution of the number of hit phototubes in the data and the simulation was observed to match very closely for central runs but, as the source position moved further up the z-axis, a discrepancy appeared. The Monte Carlo distribution appeared to peak at a lower value of NHit and with a smaller spread around the peak. This supports the theory that the detector response in the neck differs between data and Monte Carlo events, which affects the sacrifice studies. Looking at the distributions of the energy uncertainties themselves, this was demonstrated in a clear difference between data and Monte Carlo events as the run position neared the neck. The distribution in the data was broader with a noticeable tail, resulting in a larger number of events being rejected by the uncertainty cuts.

A further confirmation that this was a genuine effect of the neck came from tracking the location at which light from events hit the acrylic vessel. A clear increase in the data sacrifice was observed when the light passed through or near the neck, which was not reproduced in the Monte Carlo results.

There are many variables related to the source geometry that can have an impact on the sacrifice measurements. This was demonstrated by studying the distribution of the energy uncertainty values in the data and the simulation as a function of the distance of an event from the source. A much larger tail was observed in the uncertainty distribution near the source position in the data than in the simulation, implying that the source itself was affecting the sacrifice in data. Near the neck, other effects are introduced, as discussed above, which can magnify the source effects. It is difficult to disentangle the two and so the accepted solution was to volume weight the uncertainty due to the neck and include it in the systematic uncertainties. This procedure is described below.

Volume Uncertainties

Since the effect discussed above was clearly attributable to the neck (not a radial effect such as the one seen for the high level cuts) the detector was split into slices along the z-axis. Each slice was 50cm in height and they were chosen to extend from -575cm to +575cm, purposefully including 25cm outside the fiducial volume to incorporate finite reconstruction offsets and resolution. Events reconstructing within each bin were combined across runs and the resulting position dependence of the acceptance ratio is shown in figure 6.19.
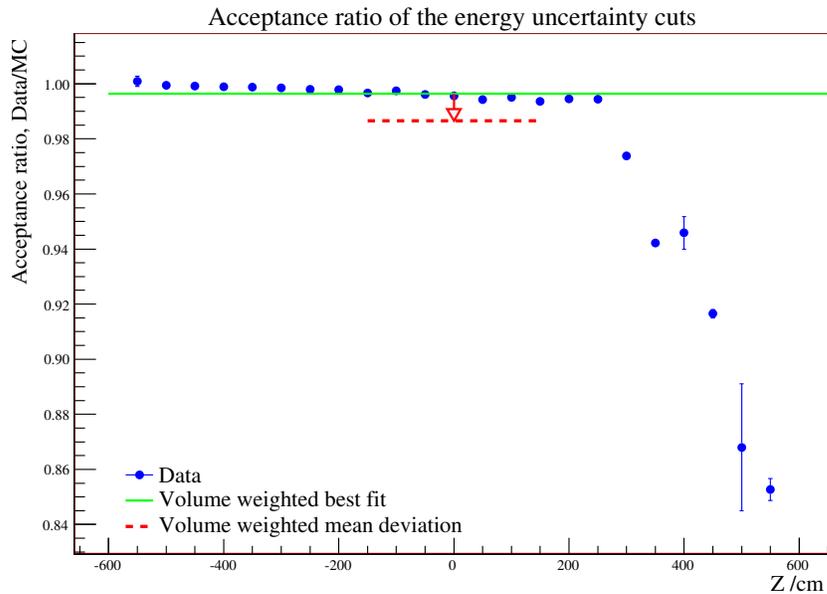
Figure 6.19: FTK uncertainty cut acceptance ratio as a function of z-position in the detector in the salt phase, using the ^{16}N source.

As for the high level cuts, the volume weighted mean across the detector was found and the volume-weighted mean deviation from this value was calculated as a measure of

the uncertainty to account for the strong tail in the distribution. However, the shape of the distribution is very clearly one-sided - at high z-positions the ratio clearly drops below 1.0, there are no points that rise above it. An asymmetric uncertainty was therefore applied. The mean deviation was used for the negative bound but, to account for the less extreme variations across the remainder of the detector, the volume weighted median deviation was calculated and applied as the positive bound. By its nature, this measure excludes the strong tail effect at the neck. This asymmetric uncertainty was added in quadrature to the other uncertainties in each energy bin.

The position uncertainties calculated in this way, using the ^{16}N source, were applied to both electron and neutron-like events. Although the ^{252}Cf source was usually used for neutron events, the source was not deployed in many positions around the detector and so would not give a reliable representation of the entire volume. In particular, the ^{252}Cf source was not deployed in the neck, which has been shown to be the most pertinent region for this particular uncertainty. However, for the positions at which the ^{252}Cf source *was* deployed, a comparison with ^{16}N shows similar results, as shown in figure 6.20. This was to be expected since the nature of the uncertainty cuts should not distinguish between the two event types for this type of measurement. If anything, the position dependence is stronger for ^{16}N than for ^{252}Cf and, so, using the uncertainty from the ^{16}N measurement was a conservative approach.
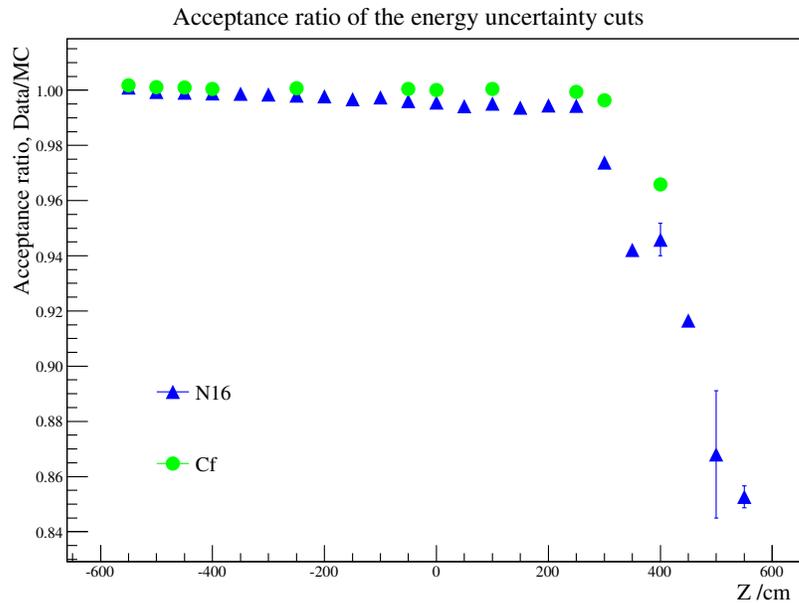
Figure 6.20: FTK uncertainty cut acceptance ratio as a function of z-position in the detector in the salt phase, comparing the ^{16}N and ^{252}Cf sources.

Spectral Uncertainties

The uncertainties relating to the energy dependence were handled differently for electrons and neutrons:

For electrons, the acceptance ratio in ^{16}N dropped very sharply above about $8\,\text{MeV}$, in a similar fashion to the result for the high level cuts, but much more extreme. Using ^{8}Li as an energy dependent check showed no such drop. Figure 6.21 illustrates this difference.
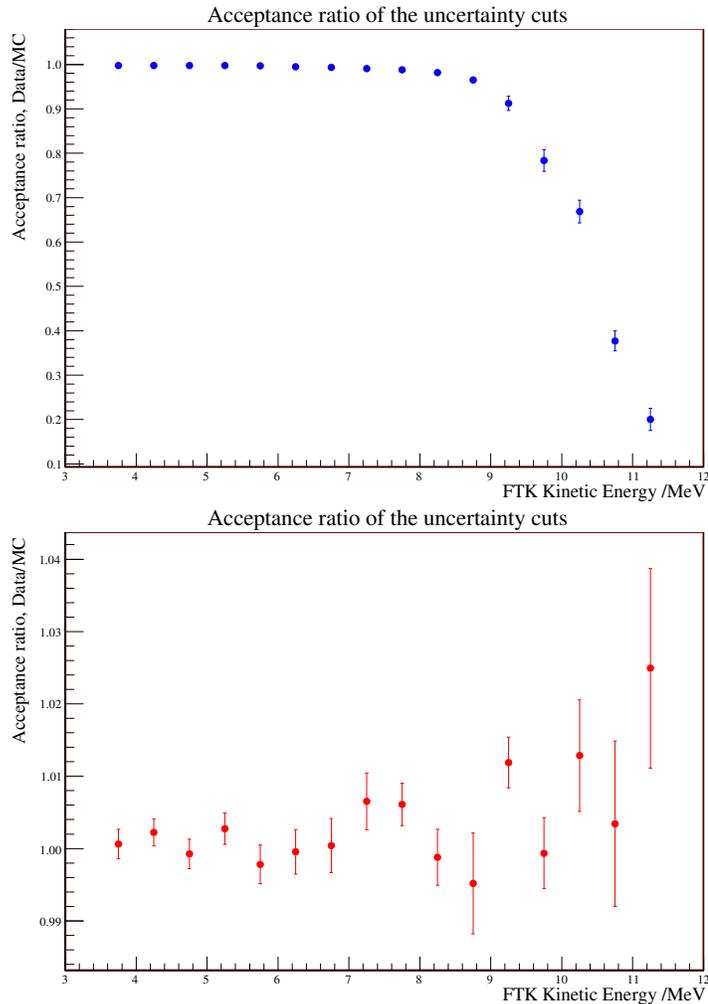
Figure 6.21: FTK uncertainty cut acceptance ratio as a function of energy in the salt phase, using the ^{16}N source (top figure) and the ^{8}Li source (bottom figure).

A check on the NHit distribution for ^{16}N did not reproduce the drop seen in the energy spectrum, suggesting that the steep drop with energy was due to low NHit ^{16}N events being misreconstructed to high energies due to source effects and, therefore, being justifiably cut by FTK. A further confirmation that the drop in the ^{16}N spectrum was

genuinely a source effect was performed using ES neutrino events. Using neutrino data for the actual measurement would not be valid, but using it to support the use of ^{8}Li over ^{16}N is acceptable. The ES events were chosen since their distinctive distribution in $\cos\theta_\odot$ (see section 3.4.1) allowed a reasonably pure sample of events to be selected. Placing an additional criterion on the energy and utilising the full sets of data cleaning cuts removed most background events from the sample. The cuts applied were as follows:

- $\cos\theta_\odot > 0.8$

- Energy $> 8.0\,\mathrm{MeV}$

- DAMN mask

- Full set of high level cuts

The sacrifice of the FTK uncertainty cuts was then calculated for both data and Monte Carlo events. The Monte Carlo result should be an accurate estimate, but the data sacrifice will be an upper limit, since some backgrounds will remain in the data sample that may also be removed by the cuts. Therefore, this will give a lower bound on the acceptance in the data and, hence, on the acceptance ratio. The lower bound found for the acceptance ratio was 0.998 using salt phase data. This was clear evidence that the drop in the ratio seen in ^{16}N was in fact a source effect, as suggested.

Given these results, in the region where the ^{16}N gives a reliable measurement, the absolute deviation of the ^{16}N bin value from 1.0 was applied as a one-sided uncertainty in that bin. Above 7 MeV, where the ^{16}N is obviously not reliable for this measurement, a two-sided uncertainty was calculated by finding the mean deviation of the ^{8}Li points from the best-fit to those points. The deviation was taken from the best-fit and not from 1.0 because, as previously mentioned, the ^{8}Li source was not necessarily trusted to give an absolute measure of the sacrifice. It was used only to verify the energy dependence.

For the neutrons, the ratio is roughly flat up to energies of about 9 MeV and then starts to rise steeply, as illustrated in figure 6.22. As seen for the high level cuts, this could be caused by source gammas occurring in coincidence with neutron events. These would be accepted in the data sample but would not occur in the Monte Carlo distribution since the source gammas are not simulated and, so, the data acceptance could rise above that of the Monte Carlo. Placing source cuts of 0.5 m and 1 m, respectively, around the ^{252}Cf source position reduced this rise from 16% to 8% and then removed it completely. The latter scenario is also shown in figure 6.22. This is strong evidence that the rise was in fact somehow linked to the presence of source gammas in the data sample. Therefore, it was assumed that this feature was not typical of neutron events. For events up to

energies of 9 MeV, the absolute difference of the bin value from 1.0 was applied as a one-sided uncertainty. Above 9 MeV the statistics are very poor for the ^{252}Cf source and, so, the value for the 8.5–9 MeV bin was used for all higher energy bins. In fact, this was a conservative choice since background contamination should fall at higher energies.
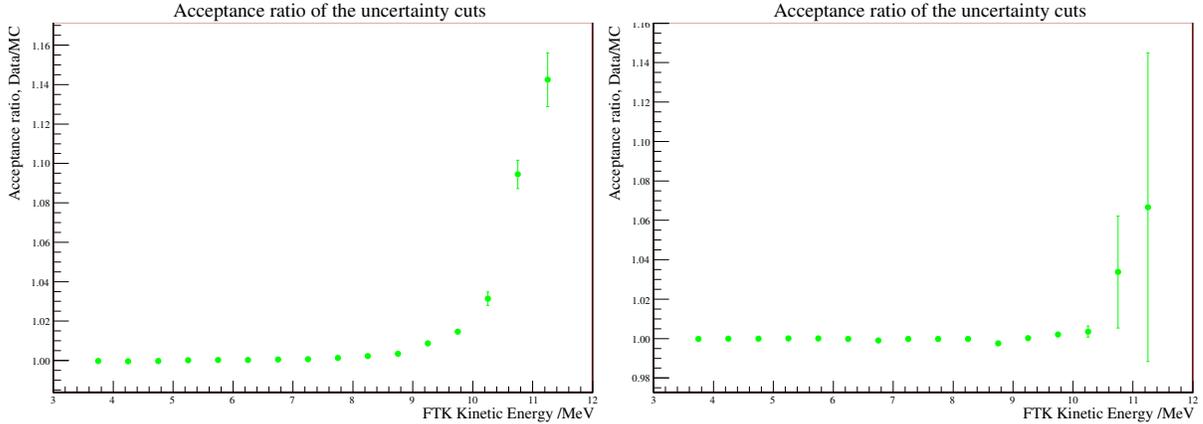
Figure 6.22: FTK uncertainty cut acceptance ratio as a function of energy in the salt phase, using the ^{252}Cf source, with no additional source cuts (left) and with an additional radial cut at 1 m around the source (right).

6.7.3 Results

In a similar manner to the high level cuts, these uncertainties were all combined in quadrature on a bin-by-bin basis. Values for each uncertainty are given in table 6.8, except the energy-related systematic uncertainties since these were calculated on a bin-by-bin basis and are best judged from figures 6.21 and 6.22. The final uncertainties were produced by combining the uncertainty on the ratio with the statistical uncertainty on the acceptance for Monte Carlo CC, ES and NC events.

6.8 Final Corrections and Uncertainties

The corrections from the four sources of sacrifice were combined to determine the final corrections to be applied to the Monte Carlo events used in the signal extraction. Since both the fitter and uncertainty cuts were assumed to have no correction, the only step was to multiply the acceptance of the instrumental cuts by the acceptance ratio of the high level cuts. This results in a global correction value to be applied to all Monte Carlo events. The uncertainties were combined in an energy dependent fashion. Within each 0.5 MeV energy bin, the uncertainties from each source of sacrifice were combined in quadrature.

Systematic Uncertainty	D$_2$O		Salt	
/%	Electron	Neutron	Electron	Neutron
Statistical uncertainty	0.0377	0.0377	0.0668	0.0322
Energy dependence	E-dep	E-dep	E-dep	E-dep
Position dependence (+)	0.0750	0.0750	0.0838	0.0838
Position dependence (−)	1.076	1.076	0.9897	0.9897
Stability	0.0834	0.0834	0.0531	0.0531

Table 6.8: FTK uncertainty cut sacrifice uncertainties, both statistical and systematic, for electron and neutron events. The (+) and (−) position dependent uncertainties refer to the positive and negative sides of this asymmetric uncertainty.

Any uncertainties that were constant across the energy range were applied with the same value in each bin.

The result is a spectrum of corrections and uncertainties (with the correction value being the same in each bin, by design) for each of the three signal types (CC, ES and NC), for each fitter, in each of the two phases of the experiment. Examples are shown in figure 6.23.

6.9 Summary

The application of data cleaning cuts in the LETA analysis results in some level of sacrifice, where signal events are removed from the data set along with the background events. A measurement of the sacrifice is a vital part of the final analysis since any uncertainties on the sacrifice translate directly into uncertainties on the extracted neutrino fluxes. Since the Monte Carlo distributions were used to define what was expected in the LETA analysis window, any factors affecting the number of events observed in that window must be fully understood and differences in the level of sacrifice between the data and simulation must be taken into account.

This chapter has described a thorough analysis of all the potential sources of sacrifice of neutrino events. Previous studies have been extended down to the lower energy threshold and new methods have been developed and applied in order to better understand the sacrifice of the new high level cuts implemented for the LETA analysis. Although these cuts have increased the overall sacrifice in comparison to previous analyses, the work presented here has measured this sacrifice with small overall uncertainties, allowing for a high level of accuracy in the extraction of the final neutrino fluxes.

The accuracy with which the energy dependence of the sacrifice has been measured has a direct impact on the extraction of a neutrino spectrum, since any differential change
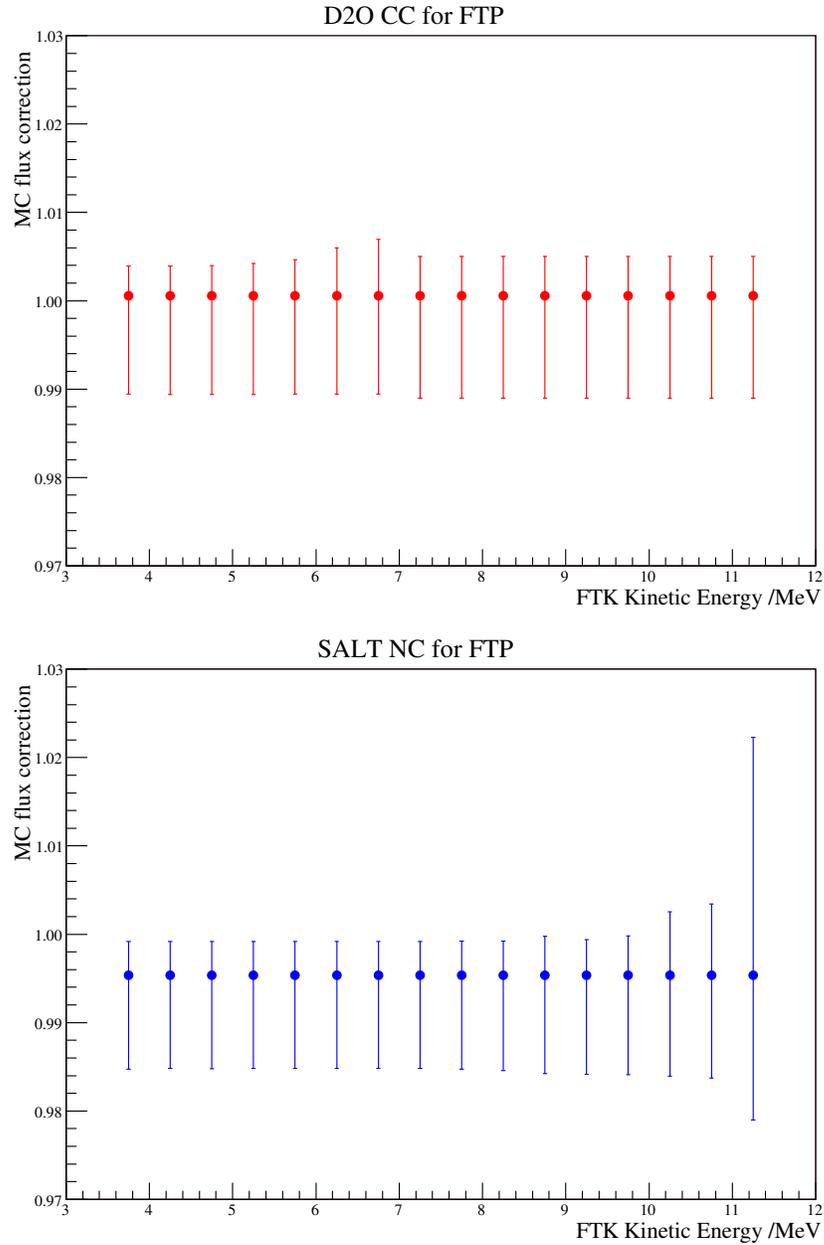
Figure 6.23: The final combined correction and uncertainties for the CC signal in the D_2O phase and the NC signal in the salt phase, using the FTK/FTP fitters.

in the sacrifice would affect the extracted signal flux differently at different energies, distorting the resulting spectrum. Therefore, particular focus was given to a differential measurement of each source of sacrifice in order to allow an unbiased extraction of the neutrino spectrum from the data set, with small associated uncertainties.

Chapter 7

Signal Extraction using the LETA Data Set

Events in the LETA data set cannot be identified as being of a particular signal type on an event-by-event basis. Instead, a set of observables was used to distinguish the signals in a maximum likelihood fit. The maximum likelihood technique is uniformly the most powerful method of parameter estimation. By maximising a product of probabilities, known as the *likelihood function*, the most probable fraction of each signal type present in the data can be extracted. This chapter describes the maximum likelihood technique and its application to signal extraction for the LETA analysis. Extensive verification of the tools used in the fit and the extraction procedure itself is presented.

7.1 The Maximum Likelihood Technique

As discussed in section 3.4, the observables that offered the strongest handle for signal separation in the LETA data set were the reconstructed energy (T_{eff}), the direction of the event relative to the Sun's location ($\cos\theta_\odot$), the reconstructed event position (ρ) and the isotropy of the observed light (β_{14}). The distribution of each observable for each interaction type was used to construct the likelihood function.

The likelihood of an event is defined as the probability of observing that event given the measured values of the observables (T_{eff}, $\cos\theta_\odot$, ρ and β_{14}) and a model for the event class. The likelihood function is given by the product of the probabilities for each event in the data set:

$$\mathcal{L} \;=\; \prod_{d=1}^{N_{Data}} F(\bar{x}_d) \tag{7.1}$$

where N_{Data} is the number of events in the data set, $\bar{x}_d$ are the values measured for the

143

observables for event d, and $F(\bar{x}_d)$ is the probability density function (PDF) giving the probability of measuring that set of values.

For a data set consisting of a number of different event types, the event probability can be broken down into the probability for each event type. In SNO, these event types are the different signal classes that make up the full data set. In this case, the probability of measuring a specific set of observable values, $\bar{x}_d$, becomes a linear combination of the probability of measuring those values for an event of each signal type:

$$F(\bar{x}_d) \;=\; \sum_{i=1}^{N_s} P_i\, F(\bar{x}_d \,|\, i) \tag{7.2}$$

$$=\; \sum_{i=1}^{N_s} \frac{n_i}{N_{Data}} F_i(\bar{x}_d) \tag{7.3}$$

where N_s is the total number of signal types in the fit and P_i is the probability of observing an event of signal type i, which is given by the fraction of events of that type in the full data set. $F_i(\bar{x}_d)$ is the PDF for signal type i, defining the probability of measuring an event with observable values $\bar{x}_d$ for that signal type. n_i is the number of events of signal type i in the data set and is the parameter that is varied in the fit in order to maximise the likelihood.

The number of events observed in the data set is actually Poisson distributed around the true mean of the model, μ_0. Therefore, for a given signal type, n_i represents a Poisson fluctuation about the true value, μ_i. In order to fit for the true value, the Poisson probability of observing N_{Data} events should be included in the likelihood function as follows, to give the *extended maximum likelihood function*:

$$\mathcal{L} \;=\; \mu_0^{N_{Data}} \frac{e^{-\mu_0}}{N_{Data}!} \prod_{d=1}^{N_{Data}} F(\bar{x}_d)$$

$$\mathcal{L} \;=\; \frac{e^{-\mu_0}}{N_{Data}!} \prod_{d=1}^{N_{Data}} \mu_0 F(\bar{x}_d) \tag{7.4}$$

This form of the likelihood function results in very large values for $\mathcal{L}$, so computationally it is more precise and more practical to maximise the logarithm of the function. Defining the expected number of events of type i as $\mu_i = \mu_0 \frac{n_i}{N_{Data}}$, such that $\mu_0 = \sum_{i=1}^{N_s} \mu_i$, and omitting constant terms in μ_0 and N_{Data} that only contribute to the overall offset, the log likelihood can be expressed as:

$$
\begin{aligned}
\mathbb{L} &= \log(\mathcal{L}) \\
&= \sum_{d=1}^{N_{Data}} \log\left(\sum_{i=1}^{N_s} \mu_i F_i(\bar{x}_d) \right) - \sum_{i=1}^{N_s} \mu_i
\end{aligned}
\tag{7.5}
$$

where μ_i are the expected number of events for signal i and are the parameters varied in the fit. These differ from n_i in that the sum is not constrained to equal the total number of observed events, thus allowing for Poisson fluctuations in each parameter.

7.1.1 Unconstrained Fits

In order to extract energy spectra for both the CC and ES neutrino fluxes from the data set, these spectra were not constrained in the likelihood fit. Instead, the flux in each energy bin was treated as a separate fit parameter. The probability for a single event (equation 7.2) was extended to sum over each spectral bin as well as each signal type:

$$
F(\bar{x}_d) = \sum_{i=1}^{N_s} \left(\sum_{k=1}^{N_{spec,i}} \mu_{ki} F_{ki}(\bar{x}_d) \right)
\tag{7.6}
$$

where $N_{spec,i}$ is the number of spectral bins for signal type i and μ_{ki} is the number of events in spectral bin k for signal i.

The integral CC and ES fluxes can be calculated from the results for the individual spectral bins using the full covariance matrix evaluated during the fit. The integral is simply the sum of the number of events in each spectral bin across the required energy range. The variance in this sum is given by the sum of the individual entries in the covariance matrix, which accounts fully for any correlations between the uncertainties on individual fit parameters.

It should be noted that, although this fit is referred to as 'unconstrained' for the purposes of this thesis, it is based on a mapping between measured energy and the true neutrino energy that assumes an undistorted spectrum. Thus, there is some degree of model assumption involved owing to the use of the Monte Carlo simulation to generate the PDFs for the fit. The magnitude of any predicted distortion is small and, so, this approximation is not unreasonable. However, a truly unconstrained fit would allow the neutrino energy spectrum to vary in the fit.

For fits to specific model predictions (such as the MSW prediction at specific values of the mixing parameters) the neutrino energies of the Monte Carlo events were directly reweighted in order to produce fully consistent PDFs for such studies. This is described in more detail in section 8.4.

Although it would be possible to fit for the true, underlying incident neutrino energy spectrum that best reproduced the observed data, in practice this can lead to complications in the interpretation of the resulting spectrum, owing to correlations. There is not a one-to-one mapping between the incident neutrino energy and the energy of the recoil electron and the finite detector resolution further smears out the resulting energy spectrum. Therefore, a wide range of incident neutrino energies contribute to the events reconstructing in a particular energy bin. This leads to strong bin-to-bin correlations in the neutrino energy spectrum, which complicate the fitting procedure and result in a spectrum that is visually difficult to interpret. For this reason, the spectra extracted from the fit were the spectra of the CC and ES recoil electron energies, as described above.

A constraint could be applied between the measured CC and ES fluxes, since the values of both are dependent on the incident ν_e flux and, therefore, the two are not independent. Although the ES interaction has a limited sensitivity to other flavours, this is known to a high precision. The NC result determines the total active neutrino content in the solar neutrino flux and the CC measures the constituent ν_e flux. Therefore, once the NC and CC fluxes are known, the ES flux is also determined, under the assumption that the total active flux is constant and consists of ν_e, ν_μ and ν_τ only. For the purposes of this thesis such a constraint was not applied, in order to test the most general case.

7.1.2 Incorporating Backgrounds in the Fit

As discussed in section 3.4.3, the lowering of the energy threshold for the LETA analysis caused a significantly greater number of radioactive background events to reconstruct inside the analysis window. Therefore, the distributions of these events had to be directly included in the extraction of the neutrino signals from the data set.

Background event types were included in the likelihood fit in exactly the same way as the neutrino signals, thus effectively increasing the number of signal types, N_s, present in the fit.

For the D_2O and H_2O backgrounds, *ex-situ* radioassays were performed to measure the concentrations of ^{214}Bi and ^{208}Tl in the detector (section 3.4.3). These concentrations can be converted into an expected number of events in the analysis window, which was used to constrain the relevant parameters in the likelihood fit. The constraints were taken to be Gaussian, but with asymmetric upper and lower 1σ values. The constraint thus contributed an extra term to the likelihood function of the form:

$$\delta\mathbb{L} \quad = \quad \sum_{i=1}^{N_s} \log\big(C_i(\mu_i)\big) \tag{7.7}$$

where C_i is the Gaussian probability of obtaining μ_i background events of type i. If no constraint exists on a particular event type, then $C_i(\mu_i) = 1$ and $\mathbb{L}$ remains unchanged.

7.1.3   Multiphase Fitting

To fit across two phases of data, the total likelihood function is just the sum of the individual likelihoods for each data set, using the relevant PDFs for each phase:

$$
\begin{aligned}
\mathbb{L} \; &= \; \mathbb{L}_{D2O} + \mathbb{L}_{Salt} \\
&= \; \sum_{d=1}^{N_{Data}} \log \left(\sum_{i=1}^{N_s} \mu_i F_i(\bar{x}_d) \right) - \sum_{i=1}^{N_s} \mu_i \\
&+ \; \sum_{d'=1}^{N'_{Data}} \log \left(\sum_{i=1}^{N_s} \mu'_i F'_i(\bar{x}'_{d'}) \right) - \sum_{i=1}^{N_s} \mu'_i
\end{aligned}
\tag{7.8}
$$

where the unprimed and primed variables are those for the D$_2$O and salt phases, respectively. For background signal types, there are no correlations between the phases[*] and, so, the numbers of events in each phase (μ_i and μ'_i) were allowed to vary independently.

Signal events were fit for in terms of the fraction of the predicted SSM flux. The solar neutrino flux should be constant across the two phases of data and so this fraction was used to relate the number of events occurring in each phase. This relation depends on both the detection efficiency and the livetime of each phase, where the livetime is defined as the total time of active data taking, taking into account any periods removed by the application of the instrumental cuts:

$$
\begin{aligned}
\text{D}_2\text{O} \; : \quad \mu_i \; &= \; f_i \times \mathcal{R}_i \times \tau_i \times \epsilon_i & (7.9) \\
\text{Salt} \; : \quad \mu'_i \; &= \; f_i \times \mathcal{R}'_i \times \tau'_i \times \epsilon'_i & (7.10)
\end{aligned}
$$

where: f_i is the fraction of the expected number of events of signal i occurring in the detector; $\mathcal{R}$ and $\mathcal{R}'$ are the rates of events predicted by the Standard Solar Model (SSM) in the D$_2$O and salt phases, respectively; τ and τ' are the livetimes of each phase and ϵ and ϵ' are the detection efficiencies. All known variations between the phases were taken into account in these parameters so, by assuming a constant solar neutrino flux, the value of f_i should be the same for each phase. By using f_i as the variable parameter in the fit, the number of events in each phase cannot vary independently.

---

[*]Constraints could be formulated for AV neutron, bulk AV backgrounds and PMT β-γs between the phases but this has not been incorporated into this analysis.

7.2 The Signal Extraction Code

The signal extraction presented here was carried out using a Fortran-based numerical routine installed in SNOMAN, called MXF. MXF utilises CERN's MINUIT [66] package to determine the set of fit parameters that maximises the log likelihood for a given data set. MINUIT is a minimisation routine and, so, in practice the negative log likelihood is minimised. MXF was created by G. McGregor [77] and significantly expanded and verified by J. Wilson [36] for the salt phase signal extraction [12]. More recently it was extensively modified by M. Dunford to incorporate joint-phase fitting for the LETA analysis [51].

Further modifications were made to the code for the purposes of the work presented here. In particular, a more efficient method for calculating the likelihoods devised by S. Seibert [78] was implemented. Previously, MXF looped over every event for each point in parameter space during the minimisation process in order to calculate the probability of the event being of each signal type. This involved a considerable number of repeated calculations, since the probability of an event being of signal type i, given fixed values of the four observables, does not change at each evaluation. In addition, the probabilities for two events that fall into the same bin in the four-dimensional space (the same bin in each of the four observables) will be identical. To take advantage of this, a new calculation was implemented at the start of the extraction process, immediately after the data set is loaded into memory. For each event in the data set, the probability of being associated with each signal type is found and a lookup table is filled with these values. To fine-tune the process, every subsequent event is checked and, if it falls into the same four-dimensional bin as a previous event, a counter is incremented for that row in the table instead of recording the event separately. This renders a loop over every event in the data set during the minimisation process unnecessary. Instead, one loop over the rows of the table is required to extract the probabilities for each signal and multiply by the number of events of that signal type for the specific point in parameter space. Prior to this alteration, a typical unconstrained fit across two phases, fitting for the three neutrino fluxes and all internal background events took on the order of 12 hours. The method described above improved the efficiency of the code such that the same fit finished in 5 minutes.

In addition to the efficiency improvements, the code was modified and expanded to handle systematic uncertainties in a new way, as discussed in section 7.5.2. In particular, a new technique for scanning the likelihood space in a particular systematic parameter was incorporated for parameters that were either correlated or uncorrelated between the phases, including the application of appropriate penalty terms to the likelihood to constrain the value of the parameter about a central value.

The normalisation of the NC interaction was altered to use the Monte Carlo simulation to predict the SSM flux, in the same way as for the CC and ES interactions. Previous

analyses had used theoretical predictions for the incident flux and for the neutron capture efficiency to predict the number of NC events occurring in the detector. Improvements in the measurement of the neutron capture efficiency for LETA [39] increased the accuracy of the simulation and thus allowed the handling of the interactions to be unified for this analysis. The method for handling the inclusion of background event types in the fit was also modified, to allow internal and external backgrounds to be treated independently when applying systematic uncertainties. Various other new features were also incorporated, including development of a new method for handling analytic PDFs.

7.3 Verification of Background PDFs

The PDFs used to represent background events in the signal extraction procedure were generated using the Monte Carlo simulation. To ensure the extraction was not biased, the shapes of these PDFs had to be verified using calibration sources. A comparison was made between calibration source data and the Monte Carlo simulation of that source to verify that the code was accurately modelling each type of event within the systematic uncertainties on each observable obtained from external analyses (such as that presented in chapter 5). The backgrounds are dominated by ^{214}Bi and ^{208}Tl decays, so the uranium and thorium sources, described in section 2.3.2, were used for this verification.

Both the uranium and thorium sources were deployed in the detector encapsulated in acrylic. In addition, a Delrin can was sometimes added around the source. The former runs are referred to as 'uncanned' and the latter as 'canned'. The dominant effect of the can was to absorb the Čerenkov light generated in the acrylic by βs emitted during the decay process. During processing, the Monte Carlo simulation was run without this can for all source runs in the D$_2$O volume. In the H$_2$O region, the simulation was run as just the bare decay process, with no source geometry simulated. For uncanned calibration runs, this is a negligible effect and the data and Monte Carlo events can be reliably compared. However, for the canned runs this makes a significant difference to the distributions and, so, modifications were required before the Monte Carlo simulations of these runs could be used in a reliable comparison with the data. The specific modifications and verifications of the effect on the Monte Carlo distributions are presented in appendix C.

The following section details the verification of the three AV PDFs (bulk AV ^{214}Bi and ^{208}Tl events and surface AV neutrons), the H$_2$O ^{208}Tl PDF and the PMT β-γs. Verification of the remaining PDFs was done by S. Seibert and is presented in [47]. The verification was done for the three observables used to create PDFs: T_{eff}, ρ and β_{14}. $\cos\theta_\odot$ was assumed to have a flat distribution for background signals, since the events should be uncorrelated with the Sun's position.

7.3.1   AV Thallium

To verify the PDF for thallium events in the bulk of the AV, data was used from the thorium source deployed near the AV. A single uncanned run existed in each phase, both of which showed good agreement in all three observables between data and Monte Carlo events but were very statistically limited. To make a better comparison, two canned runs were chosen at the edge of the D_2O volume. Both runs were deployed at the same location in the detector so the runs could be combined to improve statistics. Figure 7.1 shows the resulting distributions, which are in good agreement, and table 7.1 gives the χ^2 values for the fits of the simulation to the data.

A small difference was observed in the height of the neutron peak between the data and Monte Carlo energy spectra. To illustrate the difference, the energy spectra were fit with the sum of an exponential and a Gaussian to represent the Čerenkov tail with the neutron peak. This is shown in figure 7.1. The normalisations of the Gaussian and exponential parts of the function represent the fraction of neutron and electron type events in the spectrum. Approximately 14% additional neutrons were observed in the Monte Carlo distribution.

The exact position of the source in calibration runs can have a strong impact on the resulting distributions. The capture length for neutrons in the salt phase is longer than the average scattering length for γs, so shifting the source a few centimetres closer to the fiducial volume can increase the proportion of electron events observed in the analysis window. The Monte Carlo simulation can be used to estimate the effect of such a shift. Using AV ^{208}Tl events, the fraction of neutrons in the events observed at different radii from the AV was calculated. The neutron fraction inside the fiducial volume was found to be 26.2%. Shifting the analysis window by $5\,\mathrm{cm}$ changed this fraction by over 11% and shifting it by $10\,\mathrm{cm}$ changed the fraction by 30%.

The simulation of these runs was generated at the stated source position. The uncertainty on this source position for diagonal axis runs such as these is $5-10\,\mathrm{cm}$ [68], so the difference of 14% observed in the fraction of neutrons between data and Monte Carlo events is entirely consistent with the uncertainty on the position of the source.

Although neutrons have a lower β_{14} than electron events, this 14% excess of neutrons can be shown to have a negligible effect on the β_{14} distribution, so does not affect the agreement between data and Monte Carlo events for this observable. Using the Monte Carlo to work out the average β_{14} value for both neutron and electron events from AV ^{208}Tl, the observed shift in β_{14} due to this excess of neutrons would be on the order of 0.005, which would not significantly affect the agreement between the data and Monte Carlo distributions.

An approximate selection can be made of either electron or neutron-like events by
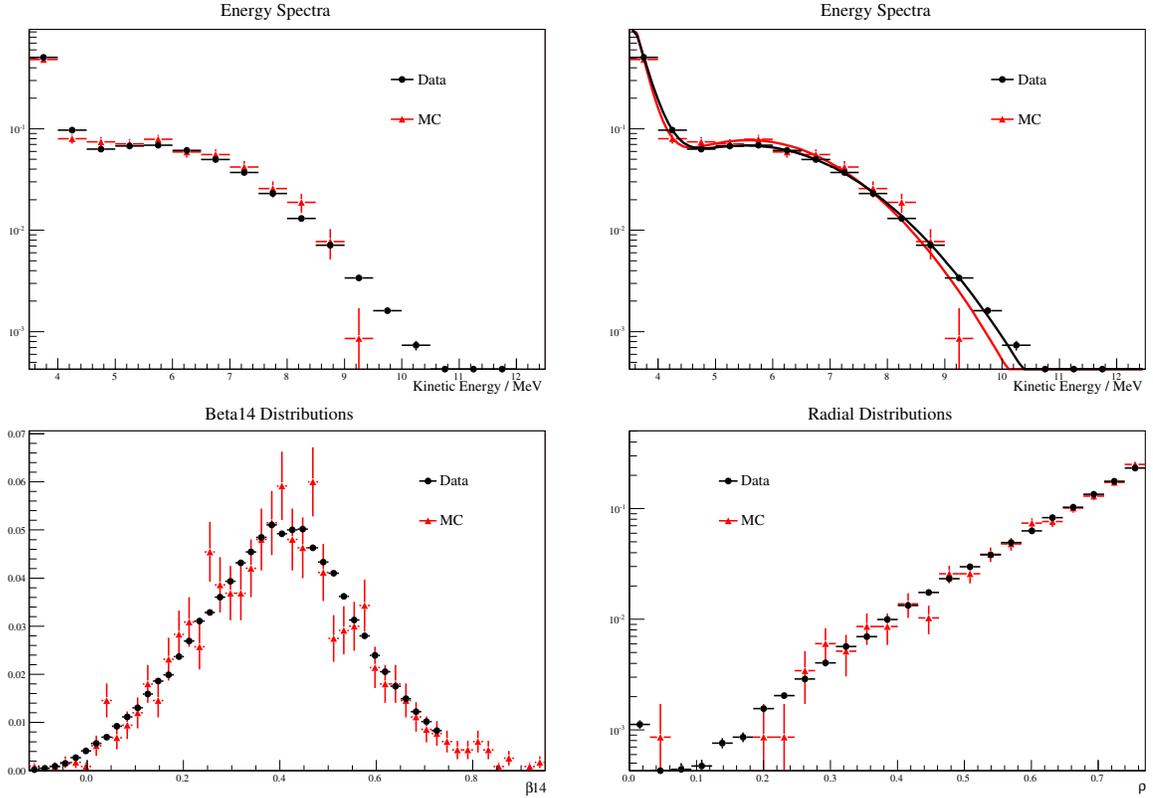
Figure 7.1: Comparison of data and Monte Carlo distributions for thorium runs near the AV in the salt phase. The energy spectra are shown both in direct comparison and with a functional form fit to the shape to illustrate the differences. The runs used were runs 25773 and 25779, with the source at (-145,0,-577).

placing a cut on both data and Monte Carlo events at 5 MeV. Events below this threshold are predominantly electron events and those above are neutrons. Using this cut, the energy spectra were compared for each type of events individually. The χ^2 values for these fits are given in table 7.1. The spectra are in clear agreement for each type of event, thus demonstrating that the discrepancy in the energy spectra was due to the normalisation differences between the two event types in data and Monte Carlo.

Taking the uncertainty due to the source positions into account, the data and Monte Carlo distributions agree well for each observable. This was taken as evidence that the thallium bulk AV PDF accurately models this source of background events. Therefore, the systematic uncertainties applied to the internal signals were also applied to these events and no further uncertainties were required to account for any discrepancies.

Observable	Distribution	χ^2/dof
Energy	Full	1.87 (12)
	Electrons only	0.59 (16)
	Neutrons only	0.88 (6)
β_{14}	Full	1.18 (47)
ρ	Full	0.82 (19)

Table 7.1: χ^2 values for the fit of the Monte Carlo distributions to the data for AV thorium runs. The χ^2 values are quoted per degree of freedom, with the number of degrees of freedom given afterwards in parentheses.

7.3.2 AV Bismuth

The bulk AV bismuth PDF was verified in an analogous manner to the thallium PDF, using the uranium source. Two canned uranium runs were chosen at the edge of the D$_2$O volume and the runs were combined to improve statistics. Taking into account the effect of the source position uncertainty on the fraction of neutron and electron events observed, as discussed above, all three distributions were in reasonable agreement for data and Monte Carlo events. The distributions are shown in figure 7.2 and the χ^2 values for the fit of the Monte Carlo events to the data are given in table 7.2.

Again, the radial and β_{14} distributions agreed very well but a difference was observed in the relative normalisations of the neutron and Čerenkov parts of the energy spectra, resulting in a poor χ^2 fit for that observable. An excess of neutrons was observed in the Monte Carlo, suggesting that the source might have been simulated too far from the fiducial volume. The source was positioned near the bottom of the D$_2$O volume for these runs, so the simulation was rerun with the source position shifted by $+3$ cm in the z direction, a shift well within the uncertainty on the source position as discussed in the previous section. The χ^2 values for fits to the data using the simulation run at both the original and shifted source position are given in table 7.2. The result of the shift in source position was to bring the energy spectra into much better agreement.

For confirmation, the χ^2 was evaluated for electron and neutron events individually using the simulation run at the original source position, employing the same cut at 5 MeV to select the event types. These values are also given in table 7.2. The energy spectrum for each event type agrees very closely with the data, illustrating that it was in fact the proportion of each event type in the distribution that caused the original discrepancy.

To confirm the expectations discussed in the previous section for the effect of a shift in source position on the proportion of neutrons observed in the spectrum, the energy spectra for the original and the shifted simulation were fit with the same functional form as previously described, consisting of an exponential plus a Gaussian. Using the
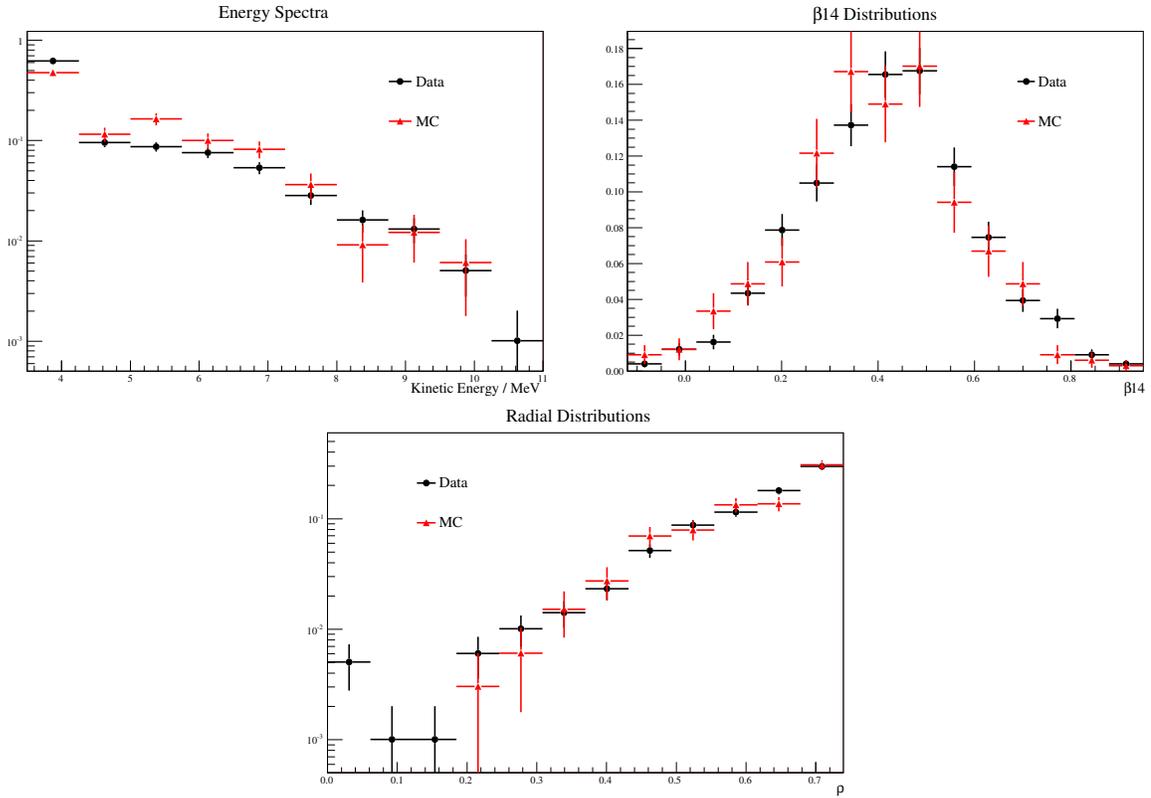
Figure 7.2: Comparison of data and Monte Carlo distributions for uranium runs near the AV in the salt phase. The runs used were 31721 and 31723.

Observable	Distribution	χ^2/dof
Energy	Full	3.42 (8)
	Electrons only	0.97 (27)
	Neutrons only	1.31 (5)
	Full, shifted MC	0.96 (10)
β_{14}	Full	1.16 (14)
ρ	Full	0.83 (8)

Table 7.2: χ^2 values for the fit of the Monte Carlo distributions to the data for AV uranium runs. The χ^2 values are quoted per degree of freedom, with the number of degrees of freedom given afterwards in parentheses.

normalisation of the Gaussian as a measure, the fraction of neutrons in the spectrum changed by 13% for the shift of 3 cm, in agreement with the predicted effect.

Bearing in mind that the shift used was well within the uncertainty on the source position of 5 – 10 cm, this was taken as evidence that the Monte Carlo accurately models the data for events of this type.

7.3.3 AV Neutrons

The ^{252}Cf source was used to model neutron events in the salt phase. Several runs were taken with the source deployed at the bottom of the D$_2$O volume, resting on the AV. These can be used to model neutron events from the surface of the acrylic. Three runs were combined and the distributions of the three observables under consideration were compared for data and Monte Carlo events detected inside the fiducial volume. The QBurst algorithm, described in section 2.3.2, was used to select neutron events from the source and an additional cut was placed to remove bursts consisting of less than three events. Such events can be caused by coincidence of two fission γs from the source and therefore contribute to contamination of the neutron sample. For runs near the AV, these events form a larger fraction of the total, since the cross-section for neutron capture on the acrylic is very high and, so, a smaller fraction of the generated neutrons are detected inside the fiducial volume.

The results are shown in figure 7.3. The χ^2 values for the fits of the Monte Carlo distributions to the data are given in table 7.3. The distributions in all three observables are in clear agreement and, therefore, no additional systematic uncertainties were required.

Observable	χ^2/dof
Energy	1.10 (30)
β_{14}	0.86 (45)
ρ	1.08 (43)

Table 7.3: χ^2 values for the fit of the Monte Carlo distributions to the data for ^{252}Cf runs near the AV. The χ^2 values are quoted per degree of freedom, with the number of degrees of freedom given afterwards in parentheses.

7.3.4 H2O Thallium

A number of uncanned thorium runs were taken in the light water region of the detector in the salt phase. Correlations are expected between observables so, for example, the energy spectra seen in the analysis window for events generated at different positions in the H$_2$O will not necessarily be in agreement. To make a fair comparison of data to Monte Carlo events, three runs at similar radii were selected for analysis and the runs were combined to improve statistics. The results are shown in figure 7.4 and the χ^2 values for the fits of Monte Carlo events to data are given in table 7.4.

The data and Monte Carlo distributions for energy and radius are in very good agreement. There is some evidence for a small offset in β_{14}. To illustrate the effect, a
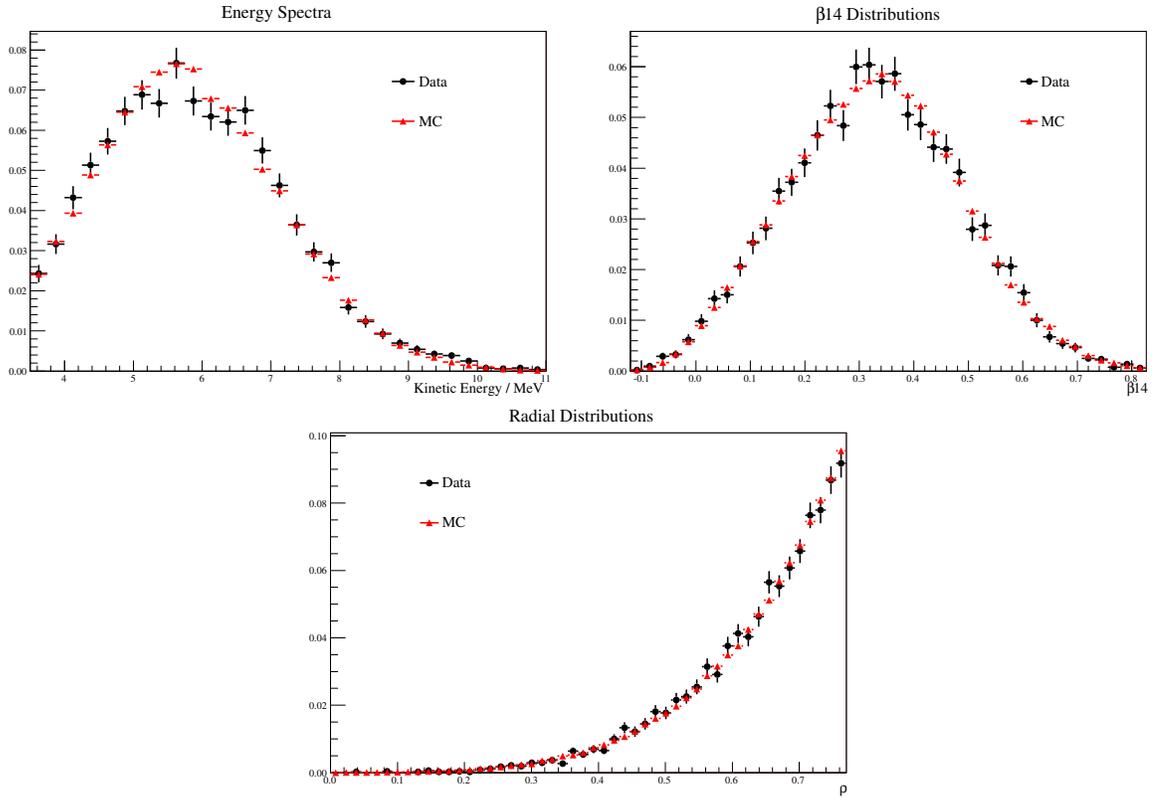
Figure 7.3: Comparison of data and Monte Carlo distributions for ^{252}Cf runs near the AV in the salt phase. The runs used were 32370-2.

Observable	χ^2/dof
Energy	1.05 (10)
β_{14}	1.23 (20)
ρ	0.80 (10)

Table 7.4: χ^2 values for the fit of the Monte Carlo distributions to the data for thorium runs in the H$_2$O. The χ^2 values are quoted per degree of freedom, with the number of degrees of freedom given afterwards in parentheses.

Gaussian was fit to each distribution. The fitted means were 0.435±0.004 and 0.418±0.006 for data and Monte Carlo, respectively, so the offset is a three sigma effect.

The runs used to produce these distributions were taken at very different z positions within the detector. It is well known that there are some residual biases in both energy and position reconstruction along the z-axis and correlations between parameters could easily cause changes in β_{14} with position. The simulation was not run at a consistent rate at each position, so combining the three runs could easily introduce a discrepancy with
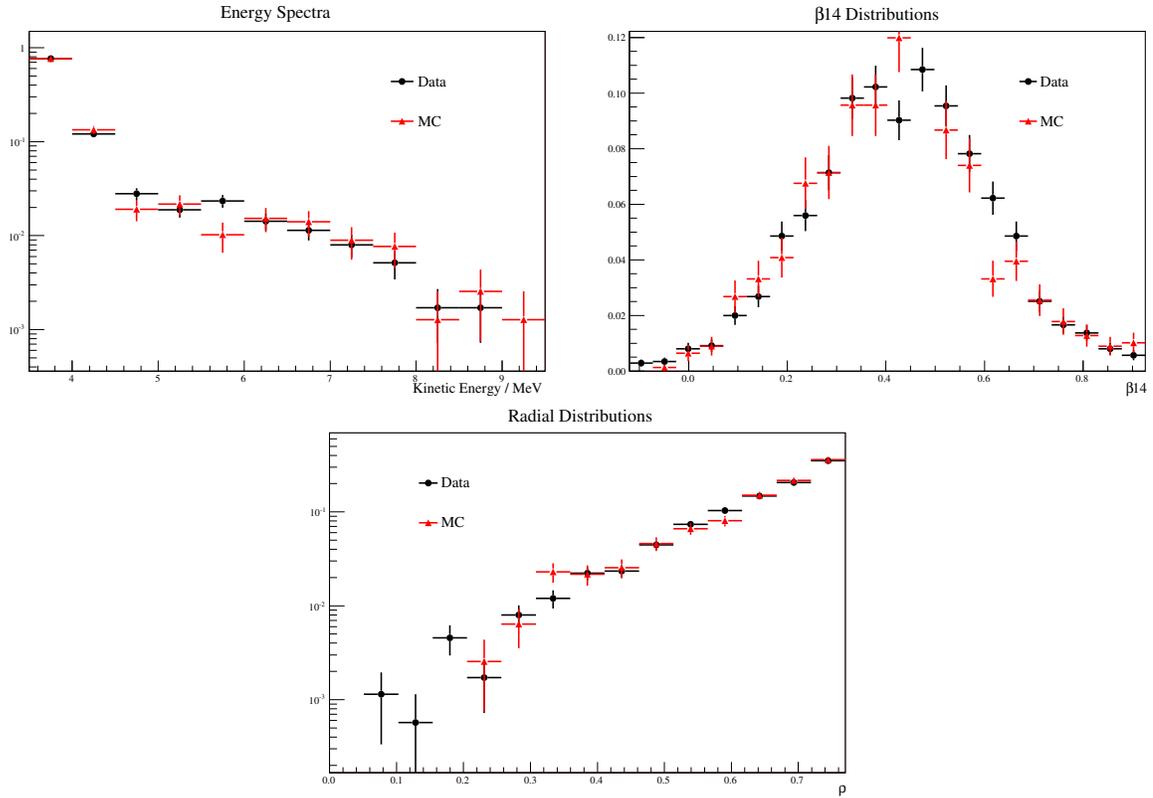
Figure 7.4: Comparison of data and Monte Carlo distributions for thorium runs in the H_2O region in the salt phase. The runs used were 24181, 29629 and 29636, all at radii of 640-660cm.

the data. To check this effect, all the H_2O-region uncanned thorium runs were recombined according to the z position of the source. The runs were separated into those positioned above and below the equatorial line. The β_{14} distributions are shown in figure 7.5 and the fit results are given in table 7.5. The data and Monte Carlo fits are in much better agreement when analysed in discrete z bins. This result implies a z-dependence to the β_{14} distribution, but a dependence that is well reproduced in the simulation. The energy and radial distributions were also rechecked for the runs split into z bins and similar agreement was seen between data and Monte Carlo events as for the runs already presented.

When analysed according to radial and z position within the detector, the data and Monte Carlo distributions for H_2O thorium runs agree to within one sigma in the statistical uncertainties for each observable. As was the case for the AV backgrounds, this was taken as evidence that the PDF for thallium decays in the light water accurately models this source of background events with no need to impose additional systematic uncertainties beyond those applied to the internal PDFs.
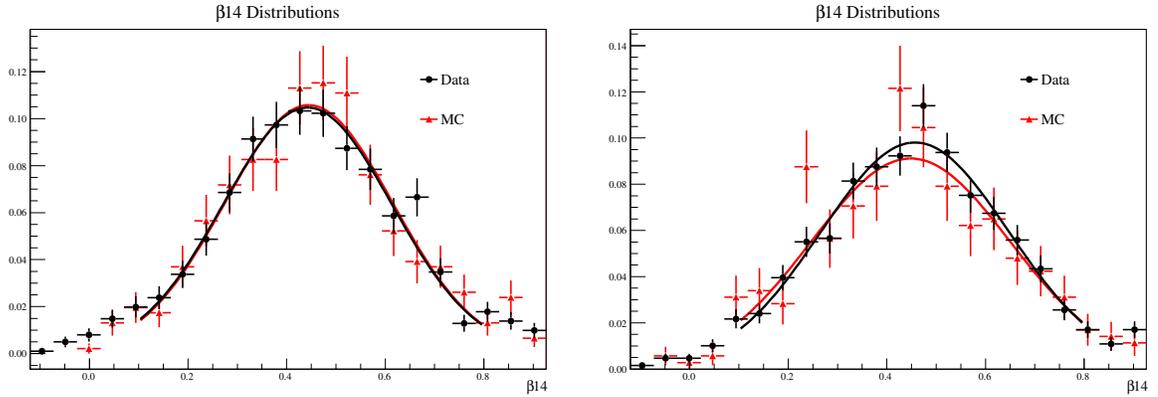
Figure 7.5: Comparison of data and Monte Carlo β_{14} distributions for thorium runs in the H$_2$O region in the salt phase. The runs used were grouped according to the source z position: above (left) and below (right) the equator ($z = 0$ cm).

Position	Event Type	χ^2/dof	Mean	Width
$z > 0$ cm	Data	0.93 (19)	0.442±0.007	0.171±0.006
	Canned MC		0.445±0.010	0.170±0.010
$z < 0$ cm	Data	0.69 (20)	0.457±0.007	0.190±0.008
	Canned MC		0.448±0.014	0.201±0.018

Table 7.5: Results of Gaussian fits to the β_{14} distributions for uncanned thorium runs in the H$_2$O region in the salt phase. The χ^2 values quoted are for the fit of the Monte Carlo distributions to the data, quoted per degree of freedom, with the number of degrees of freedom given afterwards in parentheses.

7.3.5   PMT β-γs

The PMT β-γ PDF was used to model radioactive decays originating from the PMTs, including those in the glass, the base and the steel of the support structure. External analyses demonstrated that these events were dominated by ^{208}Tl decays from thorium contamination in the materials, with little or no contribution from ^{214}Bi decays [47]. The PDF was therefore constructed from Monte Carlo simulations of ^{208}Tl decays in the PMTs.

The acceptance of PMT events is very low, since they originate so far outside the fiducial volume (the PMTs are at a radius of roughly 840 cm). However, the rate of radioactivity in the PMTs is thought to be on the order of 2 kHz [47] and so a significant number of these events are observed inside the fiducial volume due to occasional misreconstruction. No suitable calibration source exists for a sufficiently accurate verification of this PDF. The relatively low rate sources utilised in the previous sections resulted in only a handful of events reconstructing inside the fiducial volume when deployed near the

PSUP. A much higher rate thorium source was deployed in that region of the detector, but the rate was so high that the data was shown to be dominated by multiple decay events from the source coinciding within a single event window. These pile-up events had a very different topology from that of PMT events [47].

One of the high level cuts that was used to remove background events from the data set (see chapters 3 and 6) was designed specifically to target PMT events. The 'QPDT' cut looks for a single high charge tube or several hits spread across a number of PMTs that all occurred in an 'early time' window, roughly 60 ns prior to a reconstructed event. This is expected to be a typical feature of events occurring in the glass of the PMTs and so this cut could be used to tag PMT events in the data set, which could then be used to verify the shape of the PMT PDF.

A full study confirming the validity of this technique is presented in appendix D. Events that failed the QPDT cut but passed every other selection criteria applied to the LETA data set were taken to represent PMT events and a comparison was made of such events between the data set and the Monte Carlo simulation. The results are shown in figure 7.6.

Within the statistics available, the distributions appear to agree well. However, the χ^2 value for the fit of the Monte Carlo energy spectrum to data does not indicate perfect agreement and it can be seen that the simulated events have a slightly steeper distribution. Although the χ^2 value for the radial distributions indicates a good fit, it can be seen that the data systematically exhibits more events than the simulation in the tail of the distribution, near the centre of the detector.

Possible causes of these differences could be the behaviour of the QPDT cut when applied to data and Monte Carlo events, or the assumption used in the simulation that the PMT events consisted of ^{208}Tl decays only. They could also be related to any residual biases in the modelling of the concentrators in the simulation. However, without confirmation of the cause, additional systematic uncertainties were formulated to account for the differences. Approximate functional forms were fit to each distribution as a guide to determine the magnitude of these uncertainties. An exponential was fit to the energy spectra, but the shape of the distributions did not match this form exactly and so the χ^2 value of the fit of Monte Carlo events to data was used in preference to achieving perfect agreement in the fit parameters. A Gaussian was fit to β_{14} and an exponential plus a flat portion to the radial distributions, parameterised as:

$$f(\rho) = c \left(e^{a\rho} + b \right) \tag{7.11}$$

where a, b and c were free parameters in the fit. The χ^2 values for the fit of the Monte Carlo events to the data in each observable are given in table 7.6, along with the fit parameters for each distribution.

In order to bring the distributions into agreement, a positive scaling was applied to the energy of the Monte Carlo events. The result is shown in figure 7.6 and the new χ^2 value quoted in table 7.6 shows that the application of the systematic shift has brought the distributions into very good agreement. The β_{14} distributions were in good agreement and, so, the uncertainty on the fit parameters was used to represent any systematic uncertainty. The exponential portions of the radial distributions were in excellent agreement but the flat tail in the data (represented by parameter b in equation 7.11) fit to a significantly non-zero value whereas it fit to zero for the Monte Carlo events. To account for this difference, the magnitude of this tail was measured using the QPDT data events and was included in the parameterisation of the PMT PDF described in section 7.4.6. To allow for the uncertainty inherent in the difference from the simulation, the negative one sigma bound on the parameter was inflated to include zero.
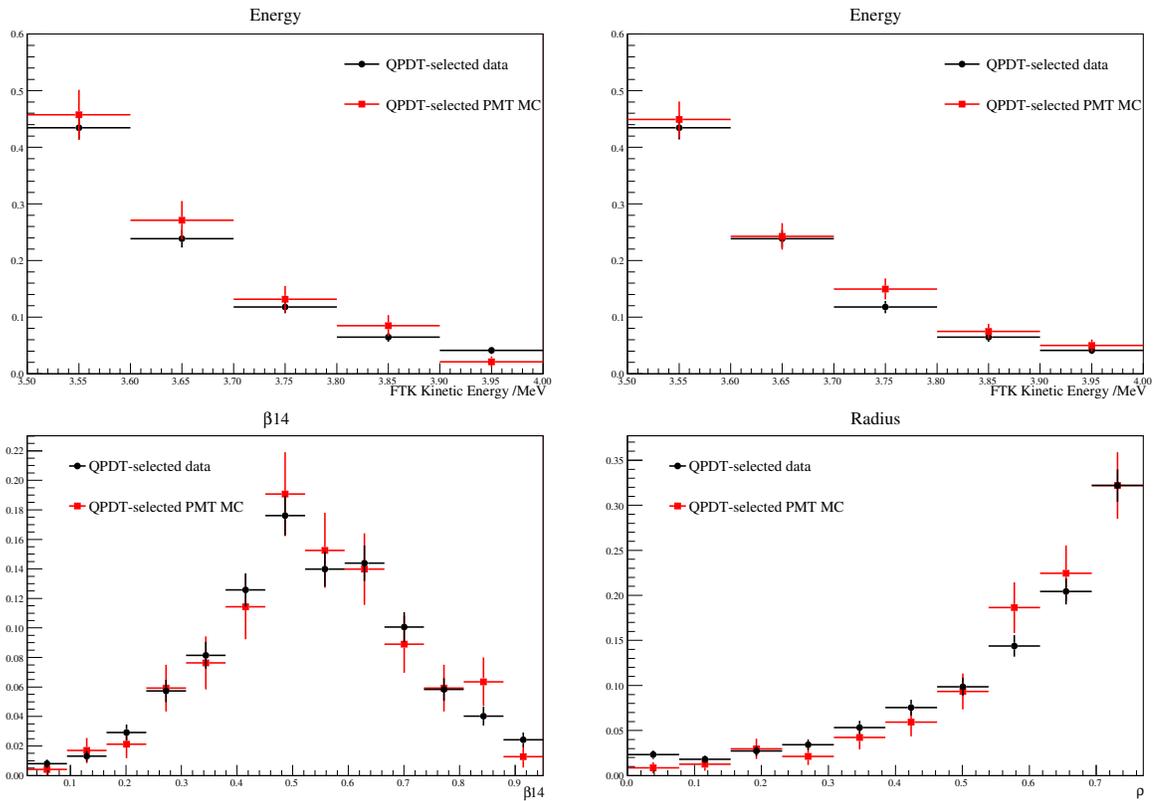
Figure 7.6: Comparison of observable distributions for QPDT-selected events in the data set and in the Monte Carlo simulation. The upper figures show the energy spectra both before (left) and after (right) applying the systematic effect described in the text to the Monte Carlo events. The lower figures show the β_{14} (left) and radial (right) distributions.

The statistics available in the D_2O phase were very low and, so, the comparison

Observable	Event Type	χ^2/dof	Fit parameters
Energy	Data		-6.051±0.002
	MC	1.32 (4)	-6.649±0.018
	Scaled MC	0.84 (4)	-6.400±0.015
Radius	Data		5.36±0.17, 1.07±0.66
	MC	1.01 (9)	5.31±0.43, -0.02±0.81
β_{14}	Data		0.534±0.010, 0.168±0.014
	MC	0.48 (12)	0.531±0.015, 0.146±0.020

Table 7.6: Results of fits to the energy, radial and β_{14} distributions for QPDT-selected events in the data and the Monte Carlo simulation, in the salt phase. The energy is given both with and without the application of the systematic scaling described in the text. The fit parameters are the exponent for energy, the exponent and flat tail for radius and the mean and width of a Gaussian for β_{14}. The χ^2 values quoted are for the fit of the Monte Carlo distributions to the data, quoted per degree of freedom, with the number of degrees of freedom given afterwards in parentheses.

presented here used the salt phase data, under the assumption that any differences between data and Monte Carlo events are expected to be due to the modelling of the specific event type in the simulation and, therefore, constant with time. The impact of the changing detector response is well modelled by the simulation and, so, should not affect the comparison. However, since the flat tail in the radial distribution was measured directly from the data, rather than from a comparison with Monte Carlo events, this was done separately for each phase to allow for the effect of the changing response.

The resulting effects that were propagated as systematic uncertainties on the shape of the PMT PDF in the signal extraction described in chapter 8 were:

- T_{eff} scaling of $+3.0\%$

- β_{14} scaling of $\pm 2.0\%$

- ρ scaling of $\pm 3.2\%$

where each effect was applied to the Monte Carlo events in an uncorrelated fashion. Additionally, a flat tail was added to the radial distribution, given by:

$$\text{D}_2\text{O:} \quad b = 3.31\,^{+1.43}_{-3.31} \qquad \text{Salt:} \quad b = 1.07\,^{+0.66}_{-1.07}$$

where b is defined in equation 7.11. The associated uncertainties were also propagated to the signal extraction. The smaller size of the tail in the salt phase, in comparison to that in the D$_2$O phase, could indicate that the effect is linked to the concentrators since the reflectivity of the concentrator petals is known to have degraded over time.

7.4 Verification of the Signal Extraction Procedure

The SNOMAN Monte Carlo simulation is a thoroughly tested and well-calibrated piece of code. It was therefore used to produce the observable distributions for each signal type from which the PDFs were created for signal extraction. Extensive verification of the shapes of these distributions for the background PDFs was presented in section 7.3.

The MXF code introduced in section 7.2 was set up to handle both analytic and binned PDFs. Although analytic PDFs can simplify the fitting process, an approximation would be involved in fitting any functional form to the observable distributions and so, to make use of the full information available about each interaction type, binned PDFs were selected for the majority of the signals in this analysis. Due to computational limitations, the Monte Carlo statistics of the PMT PDF were very poor and, so, the binned PDF created directly from these events contained significant statistical fluctuations. To smooth the shape of the PDF, a functional form was fit to the Monte Carlo distributions and this was used as the PDF for PMT β-γ events. The formulation of this model is described further in section 7.4.6.

7.4.1 Building the PDFs

Ideally, in order to retain the full information about each interaction type, PDFs would be created with dimensionality equal to the number of observable parameters. This would fully account for any correlations between the parameters. For the LETA analysis, this would involve four-dimensional histograms in T_{eff}, $\cos\theta_\odot$, ρ and β_{14} . In practice, limited statistics often makes this hard to achieve. Although biases can be introduced in the signal extraction by neglecting a correlation between two observables, distortions in the PDF shape due to statistical fluctuations can have a similar effect. It is important to preserve all the major correlations between parameters but in practice not all the parameters are strongly correlated and so a reduced dimensionality can be used.

Careful choice of the bin width also has an impact on the accuracy of the signal extraction. The bins should be narrow enough to fully define the shape of the distribution, so that no information is lost. However, if the bins are too narrow, statistical fluctuations can become a noticeable effect and distort the PDF shape.

These three factors need to be balanced when determining the factorisation and binning of the PDFs in order to maximise use of the available information while minimising any bias in the signal extraction procedure. Artificial data sets were created in order to optimise the configuration of the PDFs; this procedure is described in the following sections.

7.4.2 Artificial Data Sets

In order to investigate the effect of different PDF factorisations and binning, a number of artificial data sets were created on which the signal extraction could be tested. These sets were created from events generated using the full SNOMAN Monte Carlo simulation, in order to model the full detector behaviour as closely as possible. The set of Monte Carlo events was split in half: one half to create these 'artificial data' sets and the other to build the PDFs for these ensemble tests.

Each artificial data set was built to resemble the real data as closely as possible. Previous SNO results [12] were used to estimate the number of signal events and *ex-situ* measurements were used to estimate the number of background events expected in the analysis window.

Since the Monte Carlo events were generated on a run-by-run basis in order to simulate the full detector conditions for every run, the process of splitting these events into artificial data sets and PDFs needed to sample the full livetime of each phase. In addition, the number of events of each signal type had to be Poisson fluctuated for each data set since the extended likelihood approach implicitly assumes statistical fluctuation in event numbers. To achieve this, a random number generator was used to determine whether each event should be used for artificial data or for the PDFs. For those selected for artificial data, a second random number determined which set that event would contribute to. This approach implicitly included the required Poisson fluctuation for each signal type.

Two types of artificial data set were created: 45 sets were created containing signal and internal background events only, where the internal backgrounds were defined as those originating in the D_2O volume and the surface AV neutron events. Each of these 45 sets were fully independent. To provide a verification of the results, a second collection of 45 sets was generated by resampling from the same pool of events, thus randomising the events in each set. While each of these 45 sets was fully independent, the two collections of sets were not independent since some events were reused. The second collection of sets was intended to verify that any adjustment of the signal extraction procedure was less subject to statistical fluctuations in the particular artificial data sets.

The second type of artificial data set included the full number of background signal types. Since the Monte Carlo statistics available for the external backgrounds were limited, only 15 independent sets could be created and the number of external background events in each had to be reduced from the expected number. Table 7.7 gives the fraction of expected events that was used for each event type. The 10 % reduction in AV ^{208}Tl events results in a decrease of only 10-15 events per set due to the small number of these events expected in the data and the number of H_2O ^{214}Bi events is well constrained by *ex-situ* measurements (section 3.4.3). The reduction of these event types is, therefore,

expected to have a negligible impact on the results. The PMT events for these sets were drawn from the analytic function used to describe the shape of the PDF, defined in section 7.4.6, and so no reduction in the number of events was necessary. Again, two collections of sets were created to provide a verification of the results.

Event Type	D_2O phase fraction	Salt phase fraction
Bulk AV ^{214}Bi	1.0	1.0
Bulk AV ^{208}Tl	0.9	0.9
H_2O ^{214}Bi	0.43	0.65
H_2O ^{208}Tl	1.0	1.0
PMT β-γs	1.0	1.0

Table 7.7: The fraction of the predicted number of external background events used in the creation of artificial data sets.

7.4.3 Ensemble Testing

In order to appropriately adjust the PDF configuration and to test the signal extraction procedure, the full signal extraction was performed on each of the artificial data sets. The bias and pull, as defined in equations 7.12 and 7.13, were calculated for each set and the distributions of these values across the collection of artificial data sets were used to judge whether the extraction was bias-free.

$$Bias = \frac{N(x) - E(x)}{E(x)} \tag{7.12}$$

$$Pull = \frac{N(x) - E(x)}{\sigma(x)} \tag{7.13}$$

where $N(x)$ is the fitted number of events for event type x, $E(x)$ is the true number of that event type in the artificial data sets and $\sigma(x)$ is the statistical uncertainty on that fit parameter.

Calculated as defined above, the bias is the fractional shift in the fit result from the expected value and the pull is the significance of this shift in terms of the fit uncertainty. The bias is a test of the central value of the fit and the pull tests the fit uncertainties. For an unbiased signal extraction, both the bias and the pull should be distributed about zero. If the calculation of the statistical fit uncertainties is correct then the pull should have a width of 1.0.

7.4.4 CC and ES PDFs

In order to extract spectra from the fit, both the CC and ES signals were fit 'uncon-strained' in energy. This means that no information about the measured energy spectra was included in the fit. The flux in each energy bin was treated as a separate fit parame-ter, which means that separate PDFs were created in each bin. This implicitly accounts for correlations of the energy with every other observable, so it is just the correlations between the remaining three observables that need to be considered.

Correlations between energy and each of the remaining observables can be easily understood. Low energy events scatter more in the detector and so result in more isotropic hit patterns and less well defined event directions. This correlation of energy with both β_{14} and $\cos\theta_\odot$ can result in a second order correlation between these two parameters. For example, low energy events that are less isotropic will have better angular resolution. In addition, the reconstructed position is used to calculate both β_{14} and $\cos\theta_\odot$, so correlations with ρ might be expected.

In previous analyses ([36],[51]) the correlations of each parameter with energy were taken into account, but not the second order correlations of $\cos\theta_\odot$ with β_{14}. For the analysis presented here, the CC and ES signals were built as 3D PDFs in each energy bin, i, in order to account for all possible correlations between $\cos\theta_\odot$, ρ and β_{14} :

$$F_i(\bar{x}_d) = F_i(T_{eff}) \times F_i(\cos\theta_\odot, \rho, \beta_{14}) \tag{7.14}$$

For the purposes of constrained fits, in which the measured CC and ES energy spectra were fixed to the prediction from the SSM, the parameterisation adopted for the PDFs was:

$$F_i(\bar{x}_d) = F_i(T_{eff}, \rho, \beta_{14}) \times F_i(\cos\theta_\odot) \tag{7.15}$$

Although this neglects correlations of $\cos\theta_\odot$ with the other observables, the additional information provided by the fixed energy spectrum was shown to significantly improve the separability of the signals, thus rendering this effect negligible.

The energy resolution of the detector was roughly $1\,\mathrm{MeV}$, with some variation across the energy range of interest [47]. For this reason, the energy was binned in $0.5\,\mathrm{MeV}$ steps. For the internal background artificial data tests, the energy range considered was $3.5 – 12.5\,\mathrm{MeV}$. However, when performing the extraction for the full artificial data sets and, hence, also for the final data extraction, the number of signals included in the fit meant that MINUIT's internal limit of 50 fit parameters was reached, so the energy range was

limited to $3.5 - 11.5\,\mathrm{MeV}$. Very little information is available above $11.5\,\mathrm{MeV}$, since the neutrino flux drops steeply with energy.

For each of the other three observables, extensive testing was performed using the artificial data sets in order to optimise the width of bins used in the PDFs to balance the definition of the PDF shape against the available statistics. The use of 3D PDFs for these signals meant that the PDF shapes were more sensitive to the level of statistics available. The bin sizes were selected in order to ensure sufficient statistics across the range of each observable. The final choice for the binning and the range of each observable is given in table 7.8.

Observable	Range	Number of bins	Bin width
Energy	$3.5 - 11.5\,\mathrm{MeV}$	16	$0.5\,\mathrm{MeV}$
$\cos\theta_\odot$	$-1.0 - 1.0$	8	0.25
ρ	$0.0 - 0.77025$	5	0.15405
β_{14}	$-0.12 - 0.95$	15	0.0713

Table 7.8: Ranges and binning used for each observable for the CC and ES PDFs.

7.4.5 NC PDF

The neutron produced in the NC interaction has no memory of the neutrino energy and so the NC energy spectrum contains no information about the solar neutrino flux. The NC energy spectrum was therefore included in the fit and correlations of energy with the other observables must be taken into account in the PDF factorisation.

A strong correlation was expected between energy and β_{14} since higher energy γs scatter less often in the detector, resulting in less isotropic hit patterns. Figure 7.7 shows the correlations as predicted by the Monte Carlo simulation. This figure shows the NC energy spectrum in three β_{14} bins and the β_{14} distribution for the three lowest energy bins. The (T_{eff}, β_{14}) correlation is clearly very strong.

Some correlation was also expected for each of these observables with the reconstructed position in the detector. The effect was expected to be weaker than that observed between energy and β_{14}, since the radial profile of NC events is close to uniform (see figure 3.6). However, near the edge of the fiducial volume, where the profile dips as events occur nearer the AV, correlations can become more significant. Also shown in figure 7.7 is the energy spectrum and the β_{14} distribution for events reconstructing in the outermost three radial bins. The correlations are weaker than that observed for (T_{eff}, β_{14}) but they are non-negligible. In addition, three-way correlations are possible between the observables.
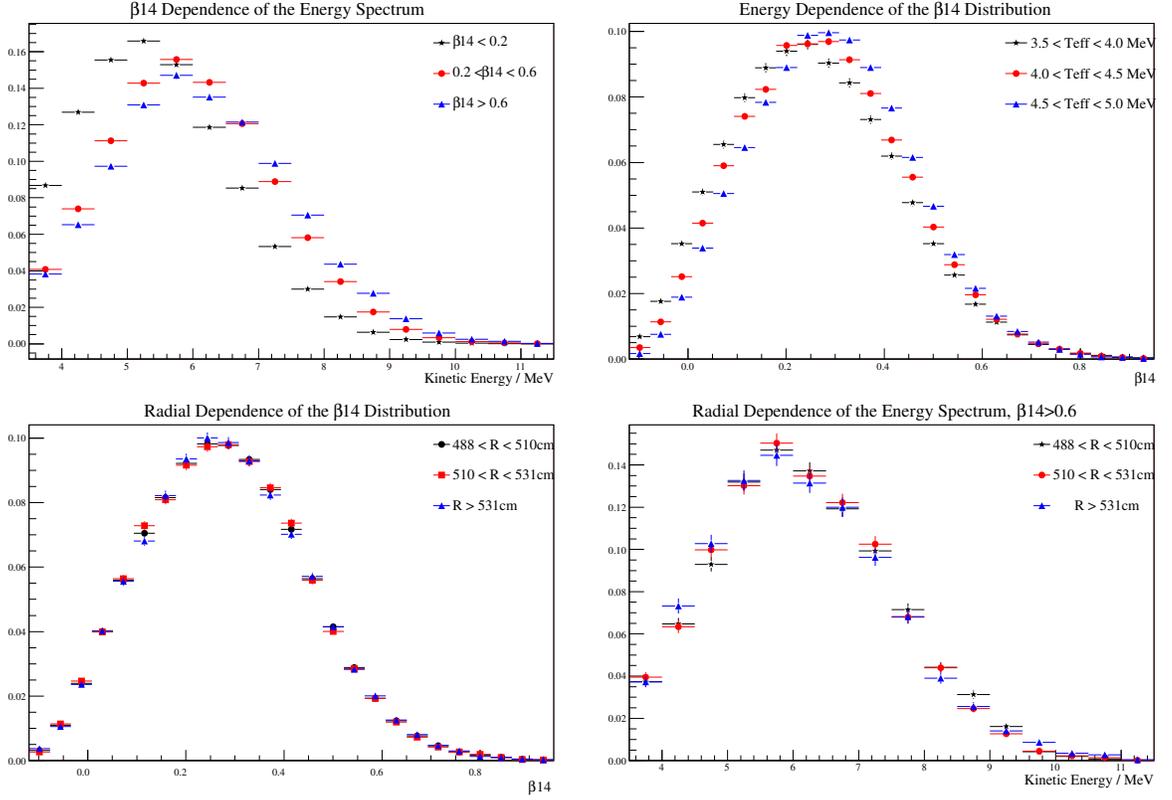
Figure 7.7: Correlations between observables for the NC signal PDF. The errors bars represent statistical uncertainties. All distributions are normalised by their area.

For this reason, all three of these observables were incorporated into a single PDF to fully account for correlations. The neutron also retains no information about the direction of the incident neutrino relative to the Sun's position and, so, the $\cos\theta_\odot$ distribution was expected to be flat for the NC signal. Therefore, the NC PDF was factorised as follows:

$$F(\bar{x}_d) = F(T_{eff}, \rho, \beta_{14}) \times F(\cos\theta_\odot) \qquad (7.16)$$

The binning used for each observable was the same as that used for the CC and ES signals, as given in table 7.8.

7.4.6 Background PDFs

Limited statistics in the background PDFs was a limiting factor in previous analyses, so for the purposes of this analysis, several extra passes of certain background event types were generated. In particular, the statistics of the internal ^{214}Bi Monte Carlo events were increased by nearly a factor of 4 over the preliminary analysis of LETA data

performed by M. Dunford [51]. This improvement significantly reduced biases observed in the extraction.

As for the NC signal, the energy spectra for the background events were included in the signal extraction. There is a strong correlation between energy and β_{14} but, due to the steep exponential nature of the energy spectra, the majority of the events fall into a small number of energy bins. For the external backgrounds, the most powerful tool for separation from the signal events was the position of the events since the distribution of these events falls off exponentially from the edge of the fiducial volume, as described in section 3.4. Therefore, the more significant parameter correlation is that between radius and β_{14}. This was shown to be particularly important in the separation of PMT events. A strong correlation was observed in the position of the peak of the β_{14} distribution with radius. In order to account for all possible correlations, the PDFs were factorised in the same way as the NC PDF (as given in equation 7.16).

Tests using the artificial data sets demonstrated that the same binning as was used for the signals resulted in some residual bias between the internal ^{214}Bi events and the low energy CC bins. Background events come from low energy radioactive decays, with a steep exponential fall off in the spectrum. In order to maximise the information available from the background energy spectra, the number of bins at the low end of the spectrum was increased to better define the spectral shape. This alteration was shown to completely remove the bias between CC and internal ^{214}Bi events at both 3.5 MeV and 4 MeV. The final binning used for the backgrounds PDFs was therefore identical to that used for the signals, as given in table 7.8, except that 0.25 MeV bins were used in the energy spectra from 3.5–5 MeV.

As mentioned above, a functional form was used to represent the PMT PDF in order to compensate for the limited Monte Carlo statistics available. The model was obtained by fitting appropriately chosen functions to each of the Monte Carlo distributions, allowing for known correlations [47]. The salt phase Monte Carlo was used in the fit, due to the higher level of available statistics. The resulting uncertainties on the fit parameters were large enough to incorporate any differences between the phases. The full functional form is given in equation 7.17. An exponential in each of energy and radius was combined with a Gaussian in β_{14}, where the mean of the Gaussian was allowed a linear dependence on ρ:

$$
\begin{aligned}
F(\bar{x}_d) &= exp(\epsilon\, T_{eff}) \times exp(\upsilon\, \rho) \times G(\beta_m, \beta_s) \qquad (7.17)\\
\beta_m &= \varpi_0 + \varpi_1 * \rho
\end{aligned}
$$

where ϵ and υ are the exponents in the energy and radial distributions, β_m and β_s are the mean and width of the β_{14} distribution and ϖ_0 and ϖ_1 are the parameters in the linear

dependence of that mean on ρ. The flat tail in the radial distribution that was observed in the QPDT study (section 7.3.5) was added to the result of the fit and is represented by b in table 7.9. Therefore, 6 parameters were used to describe the shape of the PMT PDF, the values of which are given in table 7.9. Additional systematic uncertainties were incorporated into the signal extraction to allow for the uncertainties in each parameter, the values of which are also quoted in the table. The intercept and slope of the radial dependence of the β_{14} mean had a correlation of -0.97 in the fit and so these were treated as anti-correlated in the propagation of the uncertainties in the signal extraction.

Observable	Parameter	Fit result
T_{eff}	ϵ	-7.520 $\pm$ 0.250
ρ	υ	6.059 $\pm$ 0.223
	b	D$_2$O: 3.31 $^{+1.43}_{-3.31}$, Salt: 1.07 $^{+0.66}_{-1.07}$
β_{14}	$\beta_m(\varpi_0)$	0.379 $\pm$ 0.028
	$\beta_m(\varpi_1)$	0.222 $\pm$ 0.044
	β_s	0.197 $\pm$ 0.005

Table 7.9: Parameters of the analytic function describing the shape of the PMT PDFs used in the signal extraction.

7.4.7 Ensemble Test Results

Using the final PDF configurations as described above, the full signal extraction analysis was performed on each of the artificial data sets. The results presented here focus on the energy-unconstrained fits, since the larger number of parameters varied in the fits made them more susceptible to possible biases. However, the constrained fits were also shown to be unbiased.

The results from each of the 45 internal background sets were combined and the results from the second collection of internal background sets were used as a verification. For an unbiased signal extraction, the bias and pull distribution for each fit parameter should be a Gaussian centred on zero. Figure 7.8 shows the distribution of the CC bias at 3.5 MeV and at 4 MeV. Figures 7.9 and 7.10 show the mean bias and pull for each fit parameter for one of the two collections of artificial data sets.

The distributions of biases for the low energy CC bins have been shown to be Gaussian and centred on zero, as expected for an unbiased fit. In addition, the mean bias for each of the neutrino flux parameters is consistent with zero, as shown in figure 7.9. The pulls are predominantly centred on zero with widths close to 1.0, as expected. The use of the constraint on the expected number of ^{24}Na events in the salt phase (as
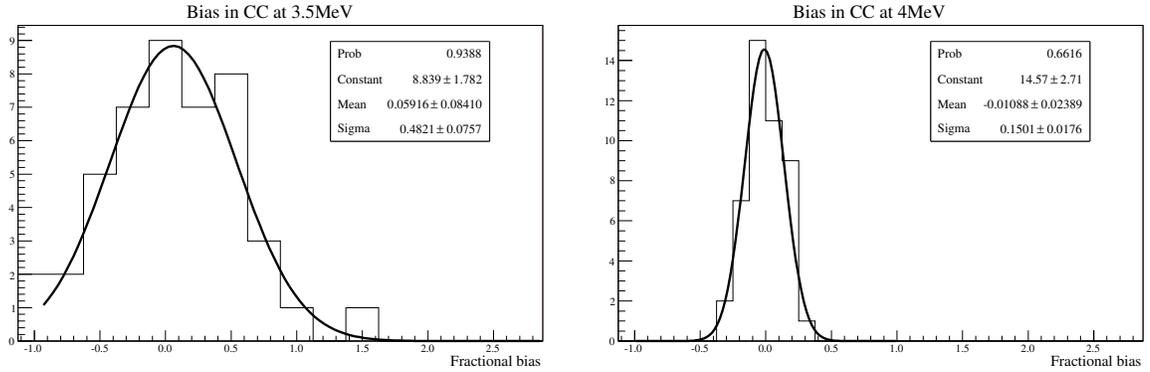
Figure 7.8: Distribution of biases for the CC flux at $3.5\,\mathrm{MeV}$ (left) and $4\,\mathrm{MeV}$ (right) from 45 internal background artificial data sets.
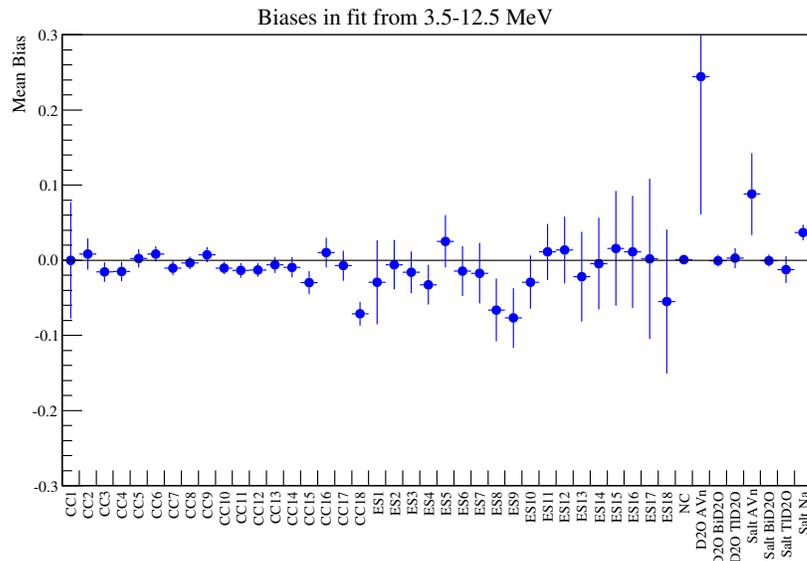
Figure 7.9: Mean fractional biases in the fit parameters from 45 internal background artificial data sets. CC1 and ES1 refer to the first energy bin of 18 spectral bins, in the range $3.5 - 12.5\,\mathrm{MeV}$.

discussed in section 3.4) narrows the width of the pull distribution for ^{24}Na and also for internal ^{208}Tl events, since there is a strong covariance between the two event types. The fact that the *ex-situ* constraints on ^{214}Bi and ^{208}Tl in the D_2O region (and similarly for the H_2O constraints on the full artificial data sets) do not narrow the pull distributions of these event types demonstrates that the fit itself is constraining the values of these parameters and the external constraint has a negligible impact.

The second collection of internal background sets resulted in very similar biases and
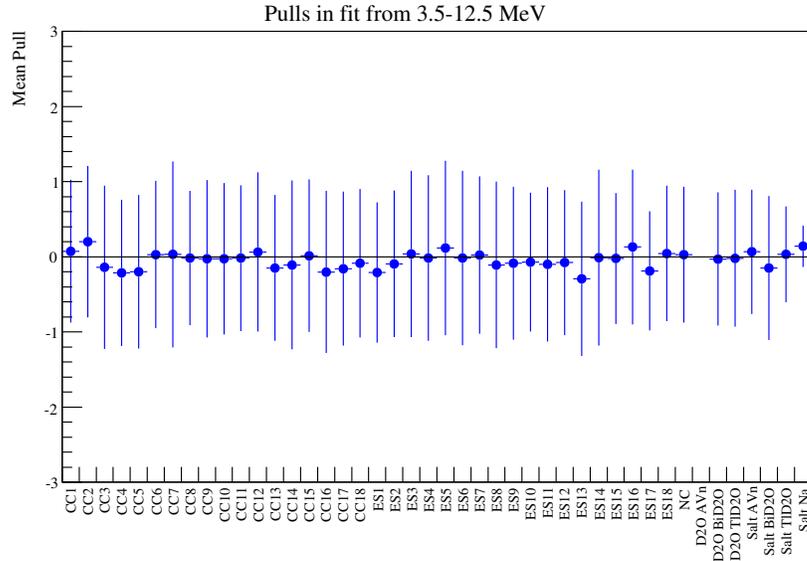
Figure 7.10: Mean pulls in the fit parameters from 45 internal background artificial data sets. CC1 and ES1 refer to the first energy bin of 18 spectral bins, in the range 3.5 – 12.5 MeV.

pulls for each of the fit parameters, demonstrating that the results were not dependent on set-to-set fluctuations and that the extraction was in fact bias-free.

One of the major challenges of the LETA signal extraction was the separation of the internal background events from the signal events. In particular, internal ^{214}Bi events consist of low energy electrons that are almost indistinguishable from CC events. Preliminary tests of the signal extraction [51] demonstrated a large bias (on the order of 40%) in the low energy CC bins due to misidentification of internal ^{214}Bi events. Both event types are uniformly distributed within the volume and have similar values for β_{14}. The $\cos\theta_\odot$ distribution was one of the few handles for separating these event types, since the CC has a weak $\cos\theta_\odot$ dependence, as discussed in section 3.4, whereas the ^{214}Bi distribution was expected to be flat since the events are generated inside the detector and, hence, are independent of the Sun's position.

As demonstrated above, this 40% bias previously observed in the CC flux at 4 MeV was completely removed by the new techniques. This bias was a major contributing factor to the limitation of that analysis to a 4 MeV threshold instead of 3.5 MeV. The inclusion of $\cos\theta_\odot$ in a three dimensional PDF for the CC signal in this analysis, thus accounting for its correlations with the other observables, played a major role in significantly reducing this bias. In addition, the optimised binning used at the low end of the background energy spectra and the extra Monte Carlo statistics generated for the background PDFs contributed to the removal of any bias in the results presented here.

The signal extraction was repeated on each of the external background artificial data sets to verify that the inclusion of AV, H_2O and PMT background events did not introduce new biases. Again, the second collection of sets was used as a verification of the results from the first and the results for the two collections of artificial data sets agreed to within one sigma statistical uncertainties. The results for both sets are shown overlaid in figure 7.11. The fit is bias-free for all the neutrino flux parameters. Some bias is observed in the background fit parameters, which is to be expected since there is not a strong handle on separating the individual background event types from each other. This is not a problem so long as the events are clearly distinguished from the signal events, which has been shown to be the case.
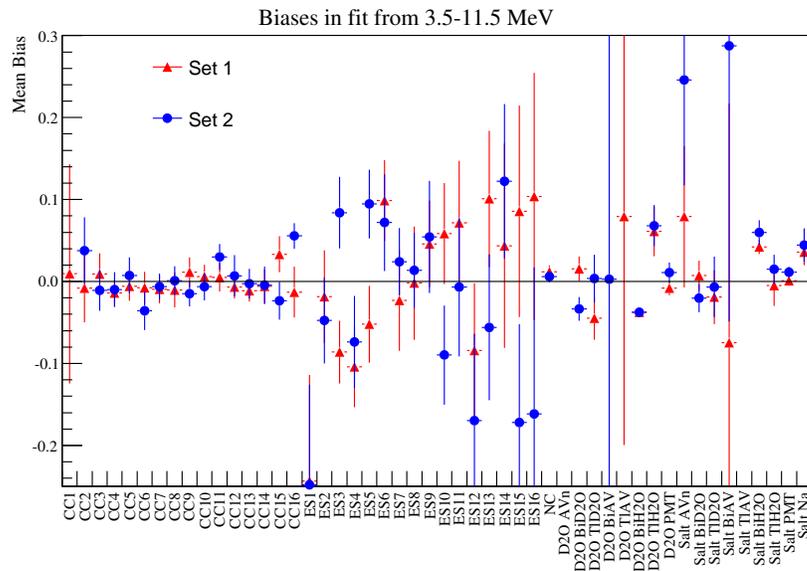
Figure 7.11: Mean fractional biases in the fit parameters from 2 collections of 15 external background artificial data sets. CC1 and ES1 refer to the first energy bin of 16 spectral bins, in the range 3.5 – 11.5 MeV.

The artificial data sets can also be used to predict the statistical fit uncertainty on the neutrino flux parameters for the final fit. Figure 7.12 shows these uncertainties for 15 external background artificial data sets. Although the uncertainty on the neutrino fluxes at 3.5 MeV is high enough that no real information is gained about the signal, the inclusion of this bin helps to normalise the number of background events, which reduces the uncertainty at 4 MeV. The statistical uncertainty observed in the 4 MeV bin in preliminary studies using a 4 MeV threshold was greater than 30%. Using the 3.5 MeV threshold in the analysis presented here has reduced this uncertainty to less than 18%.
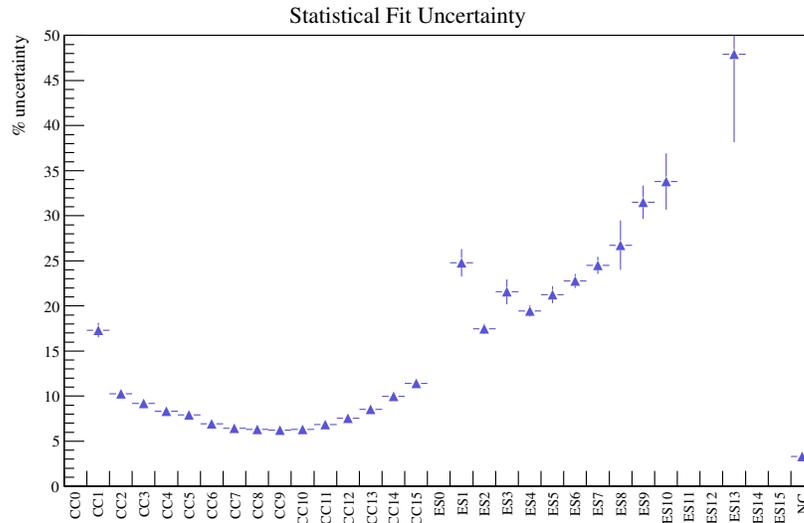
Figure 7.12: The mean percentage fit uncertainty on the neutrino flux parameters from 15 external background artificial data sets.

7.5 Handling Systematic Uncertainties

The likelihood formulation discussed in section 7.1 returns the value of each fit parameter at the best-fit point in parameter space, along with the associated statistical uncertainty. As discussed in section 3.4, the Monte Carlo simulation was used to generate the PDFs utilised in the signal extraction procedure. This means that any systematic differences between the simulation and observed data events can affect the results of the likelihood fit. Detailed studies of calibration data were used to parameterise these systematic uncertainties, such as those presented in chapters 5 and 6. The effect of these uncertainties must be incorporated into the likelihood fit and, in particular, a possible energy-dependence of these effects must be allowed for by treating the uncertainties in a differential fashion. A number of methods for dealing with these uncertainties are discussed in the following sections.

7.5.1 Modified PDFs

The accepted method for dealing with systematic uncertainties, as used in previous SNO publications, was to distort the PDFs by the measured $\pm 1\sigma$ values in each systematic parameter, redo the signal extraction to calculate the effect on the neutrino fluxes, and sum the effect of each systematic uncertainty in quadrature. The resulting systematic uncertainty on a fit parameter due to an observable systematic uncertainty, $syst$, is defined

as:

$$\delta_{syst}(x) = N'_{syst}(x) - N(x) \tag{7.18}$$

where $N(x)$ is the value of the fit parameter from the central fit (using undistorted PDFs) and $N'_{syst}(x)$ is the value from the fit using PDFs altered to account for the systematic effect, *syst*. The effects of each observable systematic uncertainty were taken to be uncorrelated with each other. The effect from each individual uncertainty was therefore summed in quadrature to give the final systematic uncertainty on each fit parameter.

When fitting across multiple phases of data, observable systematics in the two phases are not necessarily uncorrelated. The exact correlation is difficult to determine but, to first order, they can be treated as either fully correlated or fully uncorrelated between the phases. For uncorrelated uncertainties, the fit was repeated twice, with the systematic effect applied individually to each phase. For correlated uncertainties the fit was performed once, with the evaluated systematic applied to both phases simultaneously.

This *'shift-and-smear'* technique, as described above, was a conservative approach and tended to overestimate the final uncertainty, since it took no account of correlations between individual systematic uncertainties (barring the correlations between phases as mentioned above) and it also made no allowance for any information that the data might contain about the true values of the systematic parameters.

7.5.2 Use of the Data to Constrain Systematics

The most correct way to deal with systematic uncertainties on the observable parameters used in the signal extraction would be to allow the size of the systematic effects to vary within their uncertainties as parameters in the fit. For example, the energy scale could be treated as a fit parameter by allowing it to vary along with the signal fluxes and the maximum likelihood method would then return the best fit values for both the fluxes and the energy scale. Floating multiple systematic parameters in the fit would allow a fully correlated treatment of all the uncertainties and would also allow the data to help constrain the parameters to their true values.

For a signal extraction that already contains 50 free parameters, incorporating further fit parameters introduces an extra level of complexity into an already convoluted likelihood function. For a binned maximum likelihood fit such as the one presented here, allowing variation in the PDFs during the fitting procedure can result in a very choppy likelihood space. The true global minimum can be obscured by the effect of events shifting across bin boundaries, which can make it difficult to extract well defined values for the fit parameters. In addition, allowing variation of the observables used to define the PDFs

during the fitting procedure involves recomputing the PDFs for each likelihood evalua-
tion, thus significantly lengthening the processing time of the fit. Further discussion of
the complexities associated with floating systematic parameters is presented in [36].

A signal extraction allowing all systematic parameters to vary in the fit is beyond
the scope of this thesis. However, it is still possible to allow the data to constrain the
values of the key parameters. A scan of the likelihood space in a particular parameter can
be used to search for a minimum in that parameter and, when one is apparent, a shift can
be applied to the PDFs to move to that minimum. The likelihood itself can also be used
to evaluate the true uncertainties due to each systematic parameter. The construction of
the likelihood function is such that the 68% confidence level (the one sigma uncertainty
in the fit) is given by the points at which the negative log likelihood (NLL) has increased
by 0.5 from the true global minimum. This can be seen as follows:

The variance in the log likelihood function for a given fit parameter, μ_i, can be
determined from the second derivative of the function around the global maximum:

$$\sigma^2_{\mu_i} = - \left(\frac{1}{\frac{\delta^2 \mathbb{L}}{\delta \mu_i^2}} \right) \Bigg|_{\mu_i = \hat{\mu}_i} \tag{7.19}$$

where $\hat{\mu}_i$ is the value of the fit parameter μ_i at the point in parameter space that yields the
maximum value for the likelihood. MINUIT uses this method to evaluate the statistical
uncertainty on each fit parameter and to calculate the covariance matrix.

Taylor expanding the log likelihood function about the maximum in a given param-
eter gives:

$$\mathbb{L}(\mu_i) = \mathbb{L}(\hat{\mu}_i) + (\mu_i - \hat{\mu}_i) \left[\frac{\delta \mathbb{L}}{\delta \mu_i} \right]_{\mu_i = \hat{\mu}_i} + \frac{1}{2!}(\mu_i - \hat{\mu}_i)^2 \left[\frac{\delta^2 \mathbb{L}}{\delta \mu_i^2} \right]_{\mu_i = \hat{\mu}_i} + \dots \tag{7.20}$$

At the best-fit point, $\mu_i = \hat{\mu}_i$, the likelihood function is at a maximum ($\mathbb{L}(\hat{\mu}_i) = \mathbb{L}_{max}(\mu_i)$) and the first differential of the log likelihood function with respect to the fit
parameter is zero. At second order, neglecting higher order terms and using equation
7.19, the expansion reduces to:

$$\mathbb{L}(\mu_i) = \mathbb{L}_{max}(\mu_i) - \frac{(\mu_i - \hat{\mu}_i)^2}{2\sigma^2_{\mu_i}} \tag{7.21}$$

It can therefore be seen that at the point at which the fit parameter has changed
by one sigma from its best-fit value, the log likelihood function has decreased by 0.5 (and
therefore the negative log likelihood has increased by 0.5):

$$
\begin{aligned}
\mu'_i &= \hat{\mu}_i \pm \sigma_{\mu_i} \\
\mathbb{L}(\mu'_i) &= \mathbb{L}_{max}(\mu_i) - \frac{1}{2}
\end{aligned}
\qquad (7.22)
$$

7.5.3 Method

A manual 'scan' of the likelihood space in a particular systematic parameter can be achieved by performing the fit at fixed intervals in that parameter and comparing the value of the likelihood at each point. If the true minimum is significantly different from the central value, such that the change in likelihood is greater than the fluctuations due to binning effects, a new minimum can be found. In particular, more accurate values of the uncertainties can be evaluated by taking the points at which the likelihood has changed by 0.5, thus placing a further constraint on the effect of the systematic uncertainties.

For example, to scan the likelihood space in the energy scale, the energy of the Monte Carlo events used to create the PDFs (T_{eff}) was scaled according to equation 7.23 and the full fit was repeated for a range of values of the scaling parameter, δ_e.

$$
T'_{eff} = (1 + \delta_e)T_{eff} \qquad (7.23)
$$

The value of the negative log likelihood function for each fit was extracted and the likelihood space was graphed. If the data contained information about the true value of the systematic parameter in question, this would be evident in the likelihood values extracted from the scan. Taking the values of δ_e at which the likelihood was offset from the global minimum by 0.5 results in the one sigma uncertainty on that parameter and hence the one sigma effect on the flux parameters.

Separate analyses of the systematic uncertainties, such as those presented in chapters 5 and 6, provide independent measurements that can be used to constrain the values of the systematic parameters in such a scan. These measurements define the level of independent knowledge about each parameter and, so, this information can be incorporated into the formulation of the likelihood function using the same method as described in section 7.1.2 for constraining the number of background events in the fit. When scanning the likelihood space in a particular parameter, the likelihood was penalised by a Gaussian term, as in equation 7.7. Any corrections required for the systematic observables were treated separately (see section 8.2.1). Hence, the mean of the penalty term was set to zero and the width was defined by the uncertainty in the measurement of that parameter. For systematic parameters that were treated as correlated between the phases and for which

independent measurements existed in the two phases, the likelihood was constrained twice, once for each measurement.

For a parameter about which the data contains no information, this method results in the same uncertainty as the default shift-and-smear method, since the one sigma points in the likelihood are defined entirely by the penalty term. However, for some parameters the data itself constrains the true value and, so, a more accurate measurement of the one sigma uncertainty may be obtained.

7.5.4   Selection of Parameters

Scanning the likelihood space to determine the true value of and uncertainties in a systematic parameter was both a time-consuming and CPU-intensive task and so this method was not adopted for every parameter. Artificial data sets created for bias testing (section 7.4.2) were used to determine the approximate effect of each systematic uncertainty on the final extracted neutrino fluxes. For uncertainties that resulted in small effects on the final fit values, the default shift-and-smear method was used, as described in section 7.5.1.

The dominant systematic uncertainties were, in the past, those on the energy scale, β_{14} and the radial scaling [12]. The artificial data sets were used to determine which parameters would be well constrained by the data itself and, therefore, where extra information could be obtained by scanning the likelihood space in that parameter.

The investigations described here were performed with no additional penalty term on the likelihood (as introduced in section 7.5.3) in order to determine how much information was provided by the data itself before the addition of external constraints. External measurements were used purely to gauge the range across which the likelihood space should be scanned. A range of approximately $\pm 2\,\sigma$ about the measurement of each parameter was used for this investigation.

Some variation was observed across individual data sets, as might be expected due to statistical fluctuations and binning effects in the PDFs, and, so, the results are presented as the average across the 45 sets employed in the tests. Figures 7.13–7.15 show the change in the negative log likelihood from the true minimum (with unaltered PDFs) as a function of the shift applied to the systematic parameter. The average across all 45 sets is superimposed on each figure. The errors bars are representative of the spread in the likelihood across the sets at each point, but correlations between points mean these are not straightforward to interpret.

The results of these tests demonstrated that the data contains little information about the true value of systematic shifts or scalings of the radial distribution, as can be seen in figure 7.13. This shows the result of scanning the likelihood space for a scaling along the z axis in the D_2O phase, the measured uncertainty on which is $\pm 0.5\%$ (section

5.6). No significant change in likelihood was observed across the $\pm 2\,\sigma$ range in this parameter; the data adds no information beyond what has been measured externally. The systematic uncertainties on the radial PDF were therefore treated in the default shift-and-smear manner, using the externally measured values for each systematic parameter.
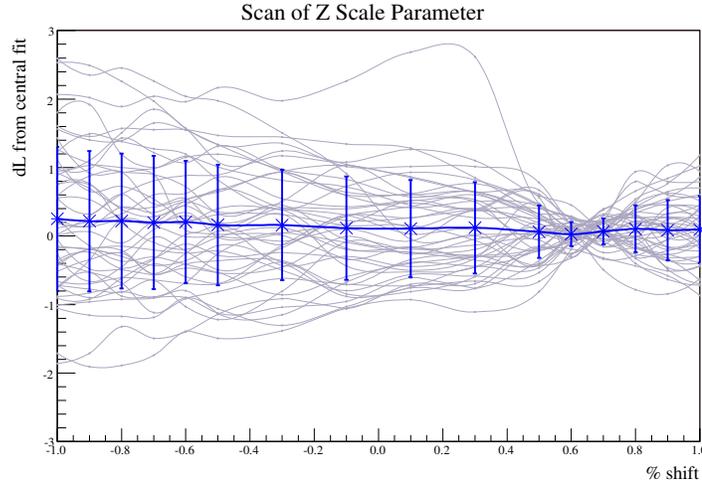
Figure 7.13: Result of scanning the likelihood space in the D_2O phase z-scale systematic parameter, using 45 artificial data sets. The result of each set is shown, with the average superimposed. The error bars represent the spread observed across the individual sets.

For the energy and β_{14} observable distributions, the data provided more information. Although individual sets were observed to fluctuate, the average forms a clear parabola around the true minimum, as illustrated in figure 7.14. This figure shows the likelihood scans for the uncorrelated energy scale in the salt phase, the uncorrelated energy-dependent fiducial volume measurement in the salt phase and the phase-correlated β_{14} scaling and width for Čerenkov events. (The correlation between phases for each of these parameters is discussed in section 8.1; in particular, the energy scale was treated as both uncorrelated and correlated, according to the source of the uncertainty). Systematic effects in both energy and β_{14} can be seen to be well constrained by the data. Therefore, scanning the likelihood space in these parameters allows the extraction of more accurate values of the uncertainties than using the externally measured constraints alone.

The likelihood space for the energy-dependent fiducial volume can be seen to be much smoother than that for the other parameters. This is because the other scans involve altering the value of one of the observable parameters used to define the PDFs, for example the energy or the β_{14} value, on an event-by-event basis (the application of each systematic effect is discussed further in section 8.1). As a result, individual events can be shifted across bin boundaries, thus affecting the likelihood evaluation in a discrete
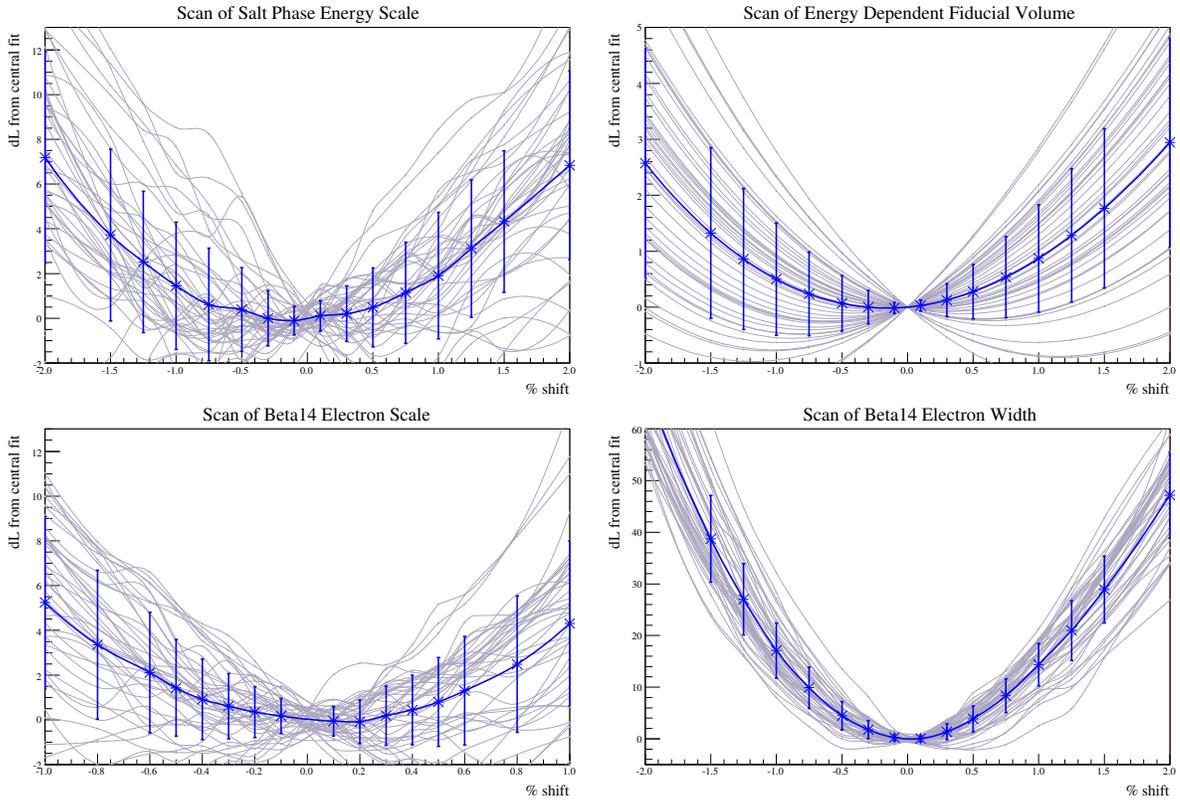
Figure 7.14: Results of scanning the likelihood space in systematic parameters, using 45 artificial data sets. The result of each set is shown, with the average superimposed. The error bars represent the spread observed across the individual sets.

manner. In comparison, the scan in the energy-dependent fiducial volume involves simply reweighting each event rather than altering the value of the observables and, therefore, individual events are not shifted across bin boundaries in the PDF. This results in a much smoother likelihood space, without the fluctuations due to binning effects.

The fluctuations seen in these tests can be used as a guide to determine the significance of any minima observed in the real data. Fluctuations on the order of a few sigma in the likelihood are common due to binning effects in the PDFs and do not represent true minima. Therefore, any minimum observed in the data must be of greater significance than this in order to be treated as the true minimum.

The investigations already discussed were performed with unshifted artificial data sets, *i.e.* where the true value of the systematic effects was known to be zero. Artificial data sets were also produced with known systematic shifts applied to determine whether the method described here could reliably extract such a shift without introducing any bias into the values of neutrino fluxes extracted from the fit.

Figure 7.15 shows an example of this. A systematic scaling of the correlated β_{14} electron mean of 0.5% was introduced into 15 full artificial data sets (containing full external backgrounds, as described in section 7.4.2) and the systematic parameter was scanned, as before. Although fluctuations in individual sets occasionally obscured the true location of the minimum, the average minimum is very close to the true value. Thus, scanning the likelihood would allow for not only more accurate determination of the uncertainties, but also the location of the true minimum.

As mentioned above, these tests were performed without the addition of the penalty term on the likelihood. The same scan performed with the penalty term applied is also shown in figure 7.15. The addition of the penalty term has the effect of smoothing the likelihood space and constraining the location of the minimum around the unshifted value (*i.e.* PDFs with no additional systematic effect applied). Although fluctuations on the order of 1–2 σ were observed in individual data sets during these tests, the addition of the penalty term smooths these and reduces the significance of any minima away from the central, unshifted point. Thus, any minima observed in the full likelihood scans with the penalty term applied that had a significance of greater than 1–2 σ were taken to be true minima and not fluctuations due to binning effects or statistics.
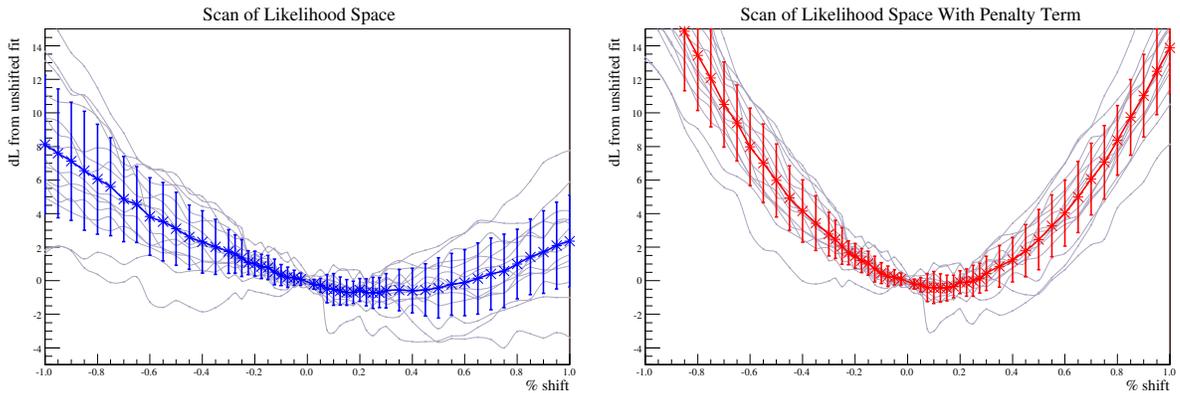
Figure 7.15: Results of scanning the likelihood space in the correlated β_{14} electron scale for artificial data sets with a systematic effect of 0.5% introduced. The result of each set is shown, with the average superimposed. The error bars represent the spread observed across the individual sets.

Figure 7.16 shows the fractional biases on the flux parameters extracted from the signal extraction at the minimum value of the systematic shift. A comparison of this figure to figure 7.11 illustrates that the use of the scan to locate the true minimum has not introduced any new bias into the extracted neutrino flux parameters.

This method for extracting minima was shown to be robust, producing accurate values for the location of the minima and unbiased neutrino flux parameters at that
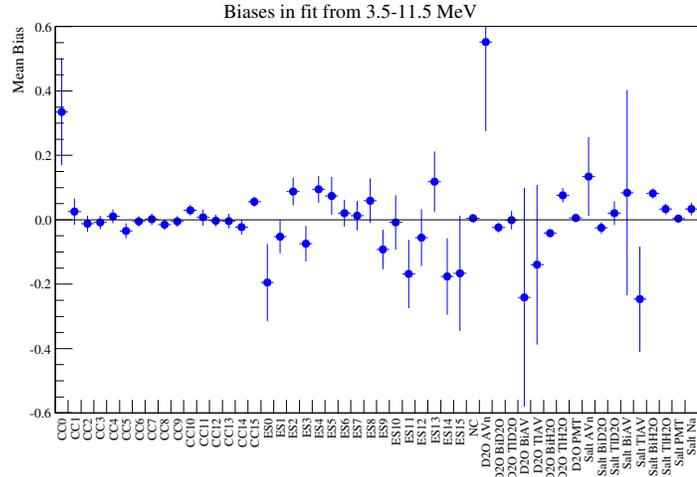
Figure 7.16: Mean fractional biases in the flux parameters for the fits at the minimum value of the systematic parameter. CC1 and ES1 refer to the first energy bin of 16 spectral bins, in the range $3.5 - 11.5\,\text{MeV}$.

point, for both the β_{14} and the energy parameters discussed above. Ideally, the parameters describing the form of the analytic PMT PDF (section 7.4.6) would also be scanned in this manner, since the PMT β-γ events were expected to be a dominant background in the fit and so obtaining the correct form for the PDF was important. However, preliminary tests into this possibility demonstrated that scanning the likelihood space had a tendency to result in biased values for the parameters. In particular, the distribution in ρ tended to be biased towards being too flat, which resulted in a positive bias in the low energy CC neutrino fluxes. At the time of writing, studies into these effects were still ongoing and, so, uncertainties in the PMT PDF parameters were treated in the default shift-and-smear method for the purposes of this thesis.

The systematic parameters that were chosen for scanning were therefore:

- The correlated energy scale

- The uncorrelated energy scale in each phase

- The uncorrelated energy resolution in each phase

- The energy-dependent fiducial volume in each phase

- The correlated β_{14} scale for Čerenkov (electron-type) events

- The correlated β_{14} width for Čerenkov (electron-type) events

Any new minima extracted from these scans were used as the central point in that parameter in the signal extraction. To allow for possible correlations between parameters, an iterative approach was used: for any new minima extracted from the scans, the entire signal extraction process was repeated, including rescanning each of the systematic parameters.

For the remaining parameters, the shift-and-smear technique was used with the measured values of the observable uncertainties provided by external analyses. The full set of systematic uncertainties applied to the data is presented in section 8.1.

To illustrate the impact of this method, figure 7.17 shows the effect of variations in some of these systematic parameters on the extracted neutrino fluxes. It can clearly be seen that the z-scaling (top left) has very little impact on the extracted flux parameters. This is to be expected, since it is the low level of its impact on the fluxes that results in such a weak constraint on the parameter from the data, as observed in figure 7.13. Therefore, the possibility of over-estimating the uncertainty by using the shift-and-smear method should not have a big impact on the final flux uncertainties. Variations in the remaining parameters have a much greater effect on the fluxes and so the more accurate values for the uncertainty resulting from the likelihood scans should yield a noticeable reduction in the final flux uncertainty in comparison to the shift-and-smear method.

7.6 Summary

This chapter has introduced the maximum likelihood method, the universally most powerful method of separating different signal types based on the values of certain observable parameters. The application of this method to the LETA analysis has been described, including the incorporation of background interactions and fitting across multiple phases of data.

The resulting signal extraction is complex, including not only the three neutrino signals but an additional 17 background PDFs defined in four different observable parameters, across two phases of data. This results in a large number of possible covariances between different signal types, particularly in the energy-unconstrained fit. In addition, the number of background events swamps the neutrino signal at the low end of the energy spectrum and, as such, small biases in the background PDFs could lead to large distortions in the extracted neutrino fluxes and spectra. Therefore, extensive testing of the procedure was required to ensure that the results were unbiased. The Monte Carlo simulation was used to generate the PDFs used for the signal extraction. The shape of the background PDFs was verified using calibration source data to ensure that no bias was introduced in to the fit via a mismodelling of one of the background event types. Ensemble tests using
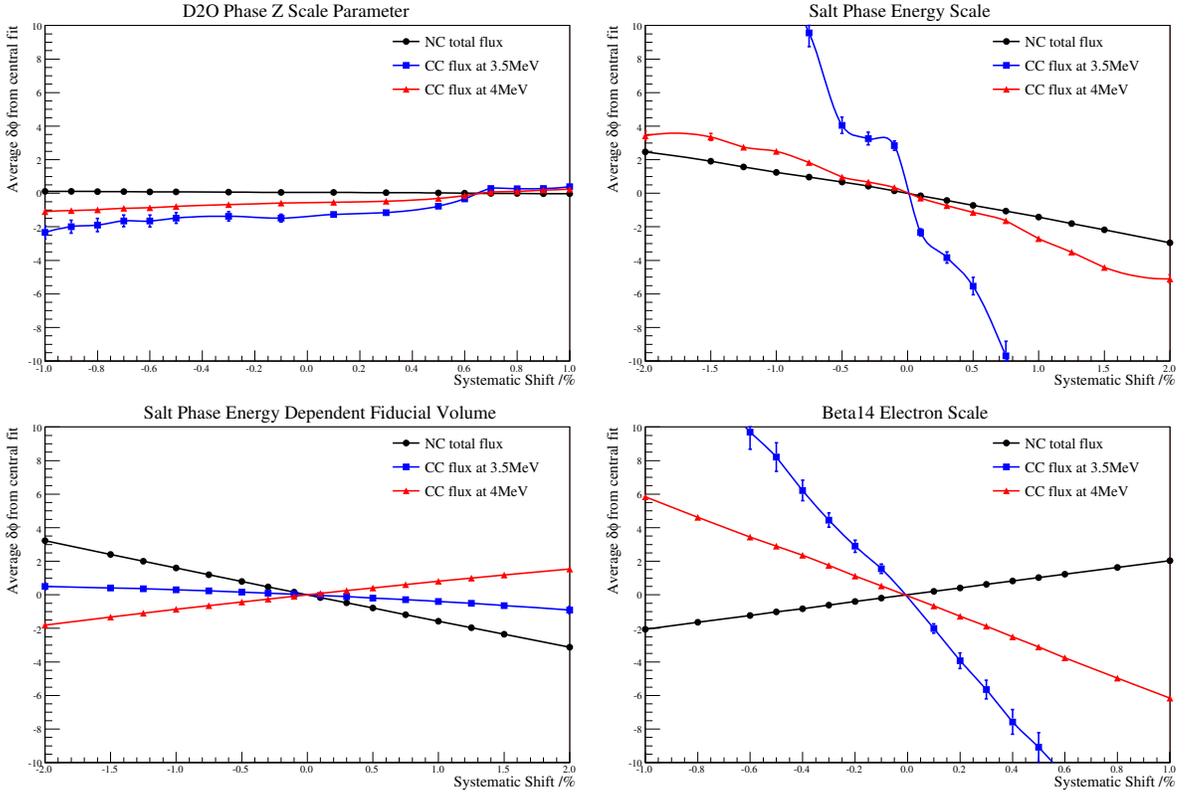
Figure 7.17: The effect of varying the systematic parameters on the extracted neutrino flux values. The average across 45 artificial data sets is shown with error bars to represent the spread observed across the individual sets.

artificial data sets demonstrated that the signal extraction was unbiased within statistics in the extraction of the neutrino flux parameters to energies as low as 3.5 MeV.

The effect of potential systematic differences in the observable parameters between data and the PDFs must be taken into account in the signal extraction. These differences can be parameterised in terms of systematic shifts or scalings applied to each event. The default method utilised by previous SNO analyses to account for these systematic uncertainties tended to overestimate the effect. An alternative technique for handling systematic uncertainties has been described that uses the likelihood evaluation to provide constraints on the value of the dominant systematic parameters. This allows for the extraction of the best-fit value for these parameters and a more accurate evaluation of the associated uncertainties, resulting in a more accurate extraction of the final neutrino fluxes and spectra.

Chapter 8

Flux and Spectrum Results

This chapter describes the extraction of the CC, ES and NC neutrino fluxes and the CC and ES energy spectra from the SNO data using an energy threshold of 3.5 MeV. Improved analysis techniques and a better knowledge of the detector, combined with the lowered energy threshold and the inclusion of two phases of data result in a more precise measurement than any previous analysis of SNO data. This is critical in determining the most likely values of the mixing parameters describing solar neutrino oscillation theory. In addition, this aids in searching for possible distortions in the spectra due to MSW mixing or new physics (see chapter 1). Interpretation of the extracted spectra requires a full understanding of the uncertainties involved. This chapter discusses the evaluation of these systematic uncertainties and presents the final extracted neutrino fluxes and energy spectra.

Also presented is a derivation of two-dimensional constraints on the values of the MSW parameters that directly utilises the likelihood value extracted from the signal extraction procedure. This allows the full information from all the observable parameters to be incorporated into the evaluation of the contours.

8.1 Systematic Uncertainties

As discussed in section 7.5, systematic differences between observed data and the Monte Carlo simulation used to generate the PDFs for signal extraction must be taken into account in the fit. Application of these uncertainties in the signal extraction falls into two categories. Systematic uncertainties on each of the observables used to create the PDFs (T_{eff}, $\cos\theta_\odot$, ρ and β_{14}) must be considered and were taken into account by distorting the shape of the PDFs by applying the systematic difference on an event-by-event basis. Additional uncertainties that affect the differential normalisation of each signal type were included by applying a weighting factor to each event. The uncertainties of each type are

discussed in the following sections.

Where possible, external analyses were performed to measure the uncertainty on each systematic parameter prior to the signal extraction being performed. For the majority of the systematic uncertainties considered here (and unless otherwise specified in the following sections) these measured values were used as the $\pm 1\,\sigma$ uncertainties in the application of the shift-and-smear technique of handling systematics described in section 7.5.1. For a number of parameters that were observed to have significant effects on the fluxes, the likelihood space was scanned in order to extract more accurate values for the uncertainties using the externally measured values as a constraint, as described in section 7.5.3.

For each observable systematic uncertainty, both correlated and uncorrelated between the phases, the fit was repeated for the positive and negative value of the uncertainty. Accounting for the full set of systematic uncertainties in this analysis therefore required the signal extraction to be performed approximately 80 times, with different modifications applied to the PDFs for each fit. The following sections describe each systematic effect that was considered, including the externally measured $\pm 1\,\sigma$ values in the observable uncertainty and the values extracted from the likelihood scan, for the parameters for which this method was employed, along with a discussion of the correlation between phases.

8.1.1 Energy

Uncertainties in the reconstructed energy of an event can be categorised into energy scale and energy resolution uncertainties. The analysis of these uncertainties can be found in [47]; the results are summarised here in order to explain their application to the signal extraction procedure.

To account for uncertainties in the energy scale, the energy of each event in the PDFs was modified as follows:

$$T'_{eff} = (1 + \delta_e)T_{eff} \tag{8.1}$$

where T_{eff} is the reconstructed event energy, δ_e is the fractional uncertainty in the energy scale and T'_{eff} is the modified energy.

A number of factors can affect the reconstructed event energy, from electronics effects to the modelling of physical processes in the detector. Those that were constant across the full span of data taking, such as uncertainties in the modelling of the ^{16}N source used to perform the measurements, were treated as correlated between the phases. Uncertainties linked to features of the detector that changed with time were treated as uncorrelated,

such as the electronics threshold and gain effects. A full discussion of each contributing factor can be found in [47]. The measured values for the correlated and uncorrelated parts of the energy scale are given in table 8.1. The measurements of the energy scale were performed using electron or γ sources, such as the ^{16}N and ^{8}Li sources. To account for possible differences in the energy scale for neutron events, an additional energy scale uncertainty was applied to neutron events only.

As discussed in section 7.5.4, the likelihood space was scanned for each of the three energy scales (correlated between the phases and uncorrelated in each phase) in order to extract more accurate values for the uncertainties on these parameters. When the scan was performed for the correlated energy scale, a minimum was observed in the constrained fit at a shift of +0.75% from the central value, with a significance of greater than 2σ. In the unconstrained fit, two minima of similar significance were observed: one at a shift of +0.75%, in agreement with that seen in the constrained fit, and a second minimum at a negative shift of -0.7%. The tests on artificial data sets presented in section 7.5.4 demonstrated that fluctuations in the likelihood due to events shifting across bin boundaries could easily cause discrete jumps in the likelihood at a level of 1–2σ. These were evident in the likelihood scan of the energy scale, making it difficult to reliably determine which of the two minima was of greater significance. Therefore, further investigation of the likelihood space around each minimum is required to determine which is the true minimum. For the purposes of this thesis, the minimum at +0.75% was chosen in order to maintain consistency between the two fit types (further discussion is given in appendix F). A correction was applied to the PDFs and the full signal extraction was performed again using this point as the new central value. The likelihood scans from before and after this shift, showing the effect of the change, are shown in appendix F. The values of the uncertainties are quoted in table 8.1 and were evaluated using the change in likelihood relative to this point. In order to account for any uncertainty inherent in the choice of minimum, the negative bound on the energy scale was inflated to -1.5% for the unconstrained fit, representing the approximate separation of the two minima, to incorporate the location of the second minimum in the one sigma bound. For the remaining uncertainties, the use of the likelihood scans to extract more accurate values has resulted in a tighter constraint on the values for both fit types.

In addition to these effects, an uncertainty in the linearity of the energy scale was included to account for any differential differences between the data and Monte Carlo distributions [47]. Differences were observed in the measured values of the effect between the two phases: a non-zero effect was measured in the D_2O phase but none was observed in salt. Possible explanations for this effect, such as multi-photon effects, would be constant with time and, so, the values from the D_2O phase were conservatively used for both phases

and this effect was treated as correlated between the phases. For the application of this uncertainty, the value for the fractional shift in equation 8.1, δ_e, was given by:

$$\delta_e = -0.0137 \times \frac{T_{eff} - 5.05}{19.0 - 5.05} \tag{8.2}$$

such that the effect was zero at the energy of the ^{16}N source, where the linear energy scale uncertainties were measured.

The energy resolution uncertainty was applied by smearing the Monte Carlo event distribution with a Gaussian of a pre-determined width, as follows:

$$T'_{eff} = T_{eff} + G(0, \zeta_e) \tag{8.3}$$

where T_{eff} is the reconstructed event energy, $G(0, \zeta_e)$ is a value randomly drawn from a Gaussian with a mean of zero and width ζ_e and T'_{eff} is the modified energy. A correction was applied to the energy resolution in the Monte Carlo simulation (section 8.2.1) and, so, the measured value of the uncertainty was given by the uncertainty on this correction. The correction was applied uniformly with energy but was measured using low energy sources such as the ^{16}N source and the radon spike. As such, the negative side of the uncertainty was applied in an energy-dependent fashion to account for the possibility of no correction being required at higher energies. The uncertainty was applied with a constant value below the energy of the ^{16}N source. Above this energy, the uncertainty was increased linearly with energy such that the magnitude of the uncertainty at the upper end of the spectrum was equal to the size of the correction, thus removing the effect of the correction at these energies. This effect was treated as uncorrelated between the phases since the resolution of the detector is known to have changed over time.

The likelihood space was scanned for the energy resolution in each phase independently and a new minimum was found at an additional smearing of 0.02 MeV and 0.025 MeV in the D$_2$O and salt phases, respectively. The full results are presented in appendix F and the extracted values for the uncertainties are given in table 8.1. In some cases, the one sigma bound on the resolution was actually found to be larger than the measured value. In particular, the upper bound on the resolution in the unconstrained fit was increased by this method. However, the use of the likelihood results in the most accurate values for the uncertainties and, so, these larger uncertainties were used in the evaluation of the final systematic uncertainty on the neutrino fluxes.

8.1.2 Isotropy

The isotropy parameter, β_{14}, is used for the separation of neutron from electron-type events. At low energies it also provides a means for distinguishing different background

Uncertainty	Measured values	Extracted values	
		Unconstrained	Constrained
Correlated energy scale (%)	+0.44 -0.95	+0.10 -1.50	+0.10 -0.10
D_2O energy scale (%)	+0.37 -0.45	+0.35 -0.05	+0.35 -0.05
Salt energy scale (%)	± 0.2	+0.10 -0.40	+0.05 -0.35
Neutron energy scale (%)	± 0.2	—	—
D_2O energy resolution (/MeV)	± 0.0313	+0.035 -0.005	+0.035 -0.008
Salt energy resolution (/MeV)	± 0.0542	+0.065 -0.015	+0.003 -0.015

Table 8.1: Systematic uncertainties in the reconstructed energy scale and resolution. The measured values are taken from external analyses [47] and the 'unconstrained' and 'constrained' values refer to those extracted from the likelihood scans for each type of fit.

types from the signals. The evaluation of uncertainties on the value of β_{14} is detailed in [47]. The results are summarised below.

As discussed in chapter 3, the isotropy of electron and salt-phase neutron events is very different. The uncertainties were therefore evaluated independently for the two event types and these were assumed to be uncorrelated. The values determined for electron events were also applied to neutron events in the D_2O phase since the event topologies are very similar, as discussed in section 3.4.1.

During the analysis of the β_{14} systematic uncertainties, several corrections to the Monte Carlo distributions were found to be necessary. Possible causes of the discrepancies are discusses in section 8.2.1 and the exact corrections applied are defined. The values of the measured uncertainties were taken as the uncertainty on these corrections, combined with additional effects such as uncertainties due to time dependence or rate effects associated with the calibration source used for the study.

The uncertainties on β_{14} were determined in terms of a fractional uncertainty on the mean of the distribution and an additional width uncertainty. To account for the former, the β_{14} of events was modified as follows, where δ_b is the fractional uncertainty in the mean:

$$\beta_{14}' = (1 + \delta_b)\beta_{14} \tag{8.4}$$

An energy dependent correction to the mean β_{14} was applied in the signal extraction and, so, the uncertainty on this correction also had to be taken into account. The correction and the associated uncertainty were evaluated relative to the energy of the ^{16}N source, which was the primary calibration source used to determine the constant scaling of the β_{14} mean in equation 8.4. This allowed the two effects to be treated as uncorrelated.

The uncertainty was applied using the following form for the fractional uncertainty, δ_b, in equation 8.4:

$$\delta_b = \delta_{be}(T_{eff} - T_{eff}^N) \tag{8.5}$$

where δ_{be} is the fractional energy dependence of the β_{14} mean and T_{eff}^N is the energy of the dominant decay branch of the ^{16}N calibration source.

The width uncertainty was handled differently for the electron and neutron uncertainties due to the different approaches used in the evaluation. The neutron width uncertainty was treated in a similar manner to the energy resolution, smearing the distribution with a Gaussian of mean zero and width ζ_b:

$$\beta_{14}' = \beta_{14} + G(0, \zeta_b) \tag{8.6}$$

For the electron width uncertainty, the distribution was expanded or contracted about the mean in an energy dependent fashion:

$$\beta_{14}' = \beta_{14} + \eta_b(\beta_{14} - \bar{\beta}_{14}) \tag{8.7}$$

where η_b is the fractional uncertainty in the width and $\bar{\beta}_{14}$ is the mean of the β_{14} distribution. This mean was calculated for each event type as a function of energy, to incorporate the known correlation between event energy and isotropy. Further details and the exact form of the functions used can be found in [47].

A correction was found to be necessary to the β_{14} mean of D_2O phase electron events. The evaluation of this correction involved reference to the salt phase uncertainties and so the uncertainties in the two phases were linked and were treated as correlated. The β_{14} electron width and the energy dependence were also treated as correlated between phases. Since the neutron β_{14} values refer to salt-phase neutron events only, the question of whether to treat these uncertainties as correlated has no meaning. The measured $\pm 1\,\sigma$ values for the β_{14} systematic uncertainties are summarised in table 8.2.

As discussed in section 7.5.4, the likelihood space was scanned for both the mean and the width of the β_{14} distribution for electron events. A minimum was observed in the β_{14} scale at a shift of 0.3%, with a significance of greater than $2\,\sigma$. The PDFs were therefore corrected to account for this and the full signal extraction was performed at this point. The full results of the scans are presented in appendix F and the extracted $\pm 1\,\sigma$ values for these uncertainties are summarised in table 8.2. The scans were performed for both the constrained and unconstrained fits and, so, results are quoted for both fit

types. The use of the likelihood has not only allowed the extraction of the true minimum, but also resulted in tighter constraints on the magnitude of these systematic effects, in particular on the value of the electron β_{14} scale.

Uncertainty	Measured values	Extracted values	
		Unconstrained	Constrained
Energy dependence (%)	$\pm\,0.0692$	—	—
Electron mean (%)	$\pm\,0.42, \pm\,0.24$	+0.025 -0.10	+0.175 -0.10
Electron width (%)	$\pm\,0.42, \pm\,0.54$	+0.225 -0.25	+0.300 -0.45
Neutron mean (%)	+0.38 -0.22	—	—
Neutron width	$\pm\,0.0045$	—	—

Table 8.2: Summary of systematic uncertainties in the β_{14} mean and width. The measured values are taken from external analyses [47]; where two values are quoted, these are independent measurements for the D_2O and salt phases, respectively. The 'unconstrained' and 'constrained' values refer to those extracted from the likelihood scans for each type of fit, for the parameters for which the likelihood space was scanned. The neutron width uncertainty is the width of the convolving Gaussian in β_{14} units.

To account for possible parameter correlations, the results for the energy scale and resolution presented above were verified at this new minimum. The likelihood scans in these parameters were rerun in order to confirm that no further new minima were observed and to ascertain the impact of the shift in β_{14} on the energy uncertainties. No additional shift was observed in the energy scale or resolution and, so, the values for the energy-related uncertainties quoted in table 8.1 were evaluated at this new minimum.

8.1.3   Reconstructed Position

The analysis of the systematic uncertainties associated with reconstructed event position was presented in detail in chapter 5. The application of the results to signal extraction is summarised here. These uncertainties have a negligible impact on the PDFs for the internal signals since their radial distribution is uniform (see section 3.4.1), but have a much greater effect on the separation of the external backgrounds due to their steep exponential tail in ρ.

The accuracy of reconstruction can differ between data and Monte Carlo events in a number of ways: a global offset can exist between the two along each of the detector axes, shifting events in a uniform way throughout the volume. A position-dependent offset would result in a scaling of event positions, effectively altering the size of the fiducial volume. In addition, the vertex resolution can differ, resulting in a broader or narrower distribution. Therefore, the uncertainties were parameterised in terms of an offset, a

scaling and a resolution for each of the three detector axes. These uncertainties were applied to the PDFs as follows:

$$x'_i = x_i + \alpha_{x_i} \tag{8.8}$$

$$x'_i = (1 + \delta_{x_i})x_i \tag{8.9}$$

$$x'_i = x_i + G(0, \zeta_{x_i}) \tag{8.10}$$

where x_i is the reconstructed position along the axis in question (x, y and z) and x'_i is the position after the modification has been applied. The offset was applied by the addition of an absolute value in centimetres, α_{x_i}, to the reconstructed position along axis x_i. The scaling was applied by multiplication of the reconstructed position by the fractional uncertainty, δ_{x_i}, for that axis. The resolution systematic uncertainty was applied by convolving the distribution with a Gaussian of width ζ_{x_i}. The latter could, by its nature, only be applied as a one-sided effect and so the impact on the extracted neutrino fluxes was assumed to be symmetric.

By construction, the evaluation of each of these three uncertainties was done independently and, hence, the three were treated as uncorrelated in the signal extraction. In addition, the analysis was designed such that the measurement of each uncertainty along each of the three axes was independent and so the uncertainties for each axis were also treated as uncorrelated. Uncertainties in the reconstructed position are dominantly due to PMT charge and timing effects. Due to the changes in the modelling of the detector across the phases, such as the time dependence to the angular response (see chapter 4), these uncertainties were treated as uncorrelated between the phases. The evaluated $\pm 1\,\sigma$ uncertainties for each systematic effect along each axis are presented in table 8.3.

In addition to these uncertainties, the accuracy of the reconstructed position can exhibit an energy dependence. Events that are higher in energy generate more Čerenkov photons and, so, a larger number of PMT hits are available for the reconstruction algorithm to utilise, resulting in greater accuracy. This essentially changes the size of the fiducial volume as a function of energy, altering the number of target deuterons available for interaction. Any difference in this effect between data and Monte Carlo events would introduce biases into the extracted neutrino spectra and hence must be taken into account. Chapter 5 described the analysis of this effect and the parameterisation of the resulting systematic uncertainty. This was applied to the PDFs by weighting each Monte Carlo event according to its energy:

$$W = 1 + \eta_\rho(T_{eff} - T^N_{eff}) \tag{8.11}$$

where W is the weighting calculated for an individual event in terms of the difference of its reconstructed energy, T_{eff}, from the central ^{16}N energy, T_{eff}^N and η_ρ is the energy dependence of the fiducial volume measurement. This weighting was defined by construction to be zero at the ^{16}N energy (~ 5.05 MeV) since the evaluation of the axial scalings was performed using the ^{16}N source and, hence, no further correction was required at this energy. This systematic uncertainty was also treated as uncorrelated between the phases since it is predominantly due to uncertainties in the modelling of the detector, which changed across the phases.

Due to the importance of this uncertainty and its impact on the fluxes (see section 7.5.4), the likelihood space was scanned for each phase in order to extract more accurate values for the uncertainty. However, the data appeared to offer no additional constraints on this effect and the results of the scan were very close to the independently measured values. Therefore, the values measured in chapter 5 were used as the measure of the uncertainty in this effect. The results of the scans are given in appendix F and the final values used for the uncertainty are given in table 8.3.

Uncertainty	D$_2$O	Salt
X offset (cm)	+1.15 -0.13	+0.62 -0.07
Y offset (cm)	+2.87 -0.17	+2.29 -0.09
Z offset (cm)	+2.58 -0.15	+3.11 -0.16
X scaling (%)	+0.08 -0.57	+0.04 -0.34
Y scaling (%)	+0.10 -0.52	+0.04 -0.26
Z scaling (%)	± 0.5	+0.07 -0.59
X resolution (cm)	3.3	3.1
Y resolution (cm)	2.2	3.4
Z resolution (cm)	1.5	5.3
Energy-dep fiducial volume (%)	+0.85 -0.49	+0.41 -0.48

Table 8.3: Summary of systematic uncertainties on the radial PDF; $\pm 1\sigma$ values taken from chapter 5. The offset is given as an absolute value in centimetres, the scaling as a percentage and the resolution uncertainty as the width of the convolving Gaussian, in centimetres. The energy-dependence to the fiducial volume measurement is given as a %.

8.1.4 Reconstructed Direction

Uncertainties in the angular resolution of the SNO detector affect the $\cos\theta_\odot$ distribution and, hence, are most relevant to the ES signal PDF. The NC and background PDFs are unaffected by uncertainties in event direction since their PDFs are flat in $\cos\theta_\odot$.

The full analysis of the angular resolution systematic uncertainty is described in [79] and the implementation of this effect was discussed in detail in [80]. The conclusion

drawn was that the effect of this uncertainty on the CC PDF was negligible and, hence, only the impact on the ES PDF needs to be taken into account. The application of this uncertainty to the ES PDF was as follows:

$$\cos\theta_\odot\,' = 1.0 + (\cos\theta_\odot\, - 1.0) \times (1.0 + \delta_c) \tag{8.12}$$

where δ_c comes from the parameterisation of the angular resolution uncertainty in [79]. The values for this parameter are given in table 8.4. Using similar reasoning as for the position uncertainties, this uncertainty was treated as uncorrelated between the phases.

Uncertainty	D_2O	Salt	Correlation
Angular resolution (%)	$\pm\,0.11$	$\pm\,0.11$	Uncorrelated

Table 8.4: Summary of systematic uncertainties on the $\cos\theta_\odot$ PDF; $\pm 1\,\sigma$ values taken from [79].

8.1.5 Neutron Detection Efficiency and Sacrifice

Additional systematic uncertainties must be incorporated to allow for the relative normalisations of the neutron and electron parts of the background PDFs and for differences in the acceptance of the data cleaning cuts for data and Monte Carlo events. The effect of these uncertainties was applied as weighting factors for the Monte Carlo events.

The neutron capture efficiency has a direct impact on the number of NC events observed in the analysis window. It also affects the detection of photodisintegration neutrons from background events and, hence, affects the ratio of neutron to Čerenkov -like events in those PDFs. To account for uncertainties in this detection efficiency, each neutron event used to make the PDFs was given a weighting of:

$$W = 1.0 + \delta_n \tag{8.13}$$

where δ_n is the fractional uncertainty in the neutron capture efficiency used in the simulation. The analysis of this uncertainty is described in [47] and the results are summarised in table 8.5. Due to the different neutron detection media in the two phases, this uncertainty was treated as uncorrelated between the phases.

The cross-section for photodisintegration has a similar effect on the relative normalisations of the neutron and electron parts of the background PDFs. Although it does not affect the NC events directly, a covariance between this interaction and the photodisintegration neutrons can lead to an impact on the NC flux. The Monte Carlo simulation

predicts the cross-section from a theoretical calculation and the uncertainty associated with this calculation must be taken into account [35]. This systematic uncertainty was applied in a similar way to the neutron capture efficiency uncertainty, but the weighting was applied only to neutron events from the background PDFs. The uncertainty on this cross-section is given in table 8.5 and was treated as correlated between the phases, since it is a theoretical uncertainty and therefore independent of detector status.

Uncertainty	D_2O	Salt	Correlation
Neutron capture efficiency (%)	$\pm$ 1.2	$\pm$ 1.2	Uncorrelated
Photodisintegration uncertainty (%)	$\pm$ 2.0	$\pm$ 2.0	Correlated

Table 8.5: Summary of neutron detection efficiency related systematic uncertainties; $\pm 1\,\sigma$ values taken from [47].

The data cleaning cuts applied to both the data and Monte Carlo events result in some level of signal loss. This is known as sacrifice and was studied in detail in chapter 6. Differences in the sacrifice for data and Monte Carlo events result in a correction that must be applied to the PDFs, discussed further in section 8.2. The uncertainty on this correction must be incorporated as a further systematic uncertainty. The evaluation of this uncertainty was performed differentially with energy, in order to account for any biases in the spectrum. The full results are presented graphically in chapter 6. The uncertainty was applied by applying a weighting to each event according to both the event type and the reconstructed energy.

8.1.6 Additional PDF Systematics

As discussed in chapter 7, the verification of the background PDFs highlighted some further uncertainties on the PMT PDF that must be taken into account in the signal extraction. In addition, the statistical uncertainty on each of the fit parameters used to define the analytic model for the PMT PDF must be taken into account. These uncertainties were combined, conservatively taking the systematic effects to be symmetric, to give the resulting uncertainties summarised in table 8.6. Due to changing detector resolution, the magnitude of these effects may alter with time and they were therefore treated as uncorrelated between the phases. The intercept and slope of the linear dependence of the β_{14} mean on ρ, ϖ_0 and ϖ_1, were treated as anti-correlated within each phase.

In addition, a further systematic was incorporated to allow for uncertainties in the modelling of the distribution of ^{24}Na events in the detector. During the central fit, this distribution was assumed to be uniform. However, it is thought that ^{24}Na may have been introduced to the detector either via the neck or the water systems and, hence, the signal

PDF	Observable	Uncertainty %
PMT β-γs	T_{eff}	$\pm$ 4.48
	ρ exponent	$\pm$ 5.44
	ρ flat tail	D_2O: $^{+43.2}_{-100}$, Salt: $^{+61.7}_{-100}$
	β_{14} mean (ϖ_0)	$\pm$ 8.37
	β_{14} mean (ϖ_1)	$\pm$ 19.9
	β_{14} width	$\pm$ 2.77

Table 8.6: Summary of systematic uncertainties on the PMT PDF; $\pm 1\sigma$ values taken from chapter 7. ϖ_0 and ϖ_1 are the intercept and slope of the linear dependence of the β_{14} mean on ρ (section 7.4.6).

extraction was performed with different radial distributions for the ^{24}Na PDF to take this into account. The modified distributions assumed that events originated either near the neck or at the bottom of the detector, with a 10% linear gradient along the z-axis.

An additional uncertainty could be incorporated into the fit to account for the theoretical uncertainty in the shape of the ^{8}B spectrum. However, this is on the order of 0.5% across the energy range in question [65] and, therefore, is negligible in comparison to other effects. Such a small distortion in the incident spectrum would be smeared out by the finite detector resolution, resulting in an insignificant effect on the extracted spectra. Theoretical uncertainties in the neutrino cross sections could also be propagated but, again, the magnitude of the uncertainties is small (<1%). Therefore, for the purposes of this thesis these were assumed to have a negligible impact on the final results.

8.2 Corrections

8.2.1 Corrections to Observable Parameters

A number of corrections were required to account for residual differences between data and Monte Carlo observables that were not accounted for in the processing.

An offset of the laserball position along the z-axis during calibration of the PMT timing resulted in an offset being introduced in the data to reconstructed positions along this axis. A shift was therefore applied to all data events to correct for this effect.

In addition, a number of corrections were applied to the reconstructed energy of each event. In the D_2O phase, the global collection efficiency was modified as a function of time in order to model time dependent changes in the detector response. For both the D_2O and salt phases, a spatial correction was devised for the FTK energy to account for variations in the energy scale across the detector volume [47]. In the D_2O phase, a further radial correction was required for the FTK energy of Monte Carlo events. Amongst other

effects, this accounted for small residual variations between the angular response model used by FTK and that used in the Monte Carlo simulation. A 'multi-photon' correction was applied to both data and Monte Carlo events in both phases in order to account for the probability of multiple photons striking the same PMT during an event. The exact functional forms of the radial, drift and multi-photon corrections are described in [51] and the spatial correction is given in [47].

A Gaussian smearing was also applied to the Monte Carlo events to account for the broader energy resolution observed in calibration data in comparison to the simulation [47]. This correction was evaluated independently for the two phases since the detector resolution changes over time.

An energy dependent correction was applied to the β_{14} of both electron and neutron events. This correction was constructed to have zero offset at the energy of the ^{16}N calibration source since the remaining β_{14} analyses were performed using that source. A correction was applied to the β_{14} of events in the D_2O phase to account for the simulation having been run with too large a value of the Rayleigh scattering length and, hence, not simulating enough scattering. Increasing the amount of scattering in the simulation directly correlates with an increase in the isotropy of the events and, hence, a decrease in the value of the isotropy parameter, β_{14}. In addition, the β_{14} of salt phase neutrons was corrected, both by a fractional offset and a resolution effect, to account for differences observed between data and Monte Carlo events during calibration source studies.

The full set of corrections applied to the observables is summarised in table 8.7.

Observable	Phase	Event type	Correction
z	D_2O & Salt	Data	$z' = z - 5.0\,\mathrm{cm}$
T_{eff}	D_2O	Data	Drift correction
	D_2O	MC	Radial correction
	D_2O & Salt	Data	Spatial correction
	D_2O & Salt	All	Multi-photon correction
	D_2O	All	$T_{eff}' = T_{eff} + G(0, 0.15)$
	Salt	All	$T_{eff}' = T_{eff} + G(0, 0.145)$
β_{14}	D_2O	MC	$\beta_{14}' = 0.9919\ \beta_{14}$
	Salt	MC neutrons	$\beta_{14}' = 0.9856\ \beta_{14}$
	Salt	MC neutrons	$\beta_{14}' = \beta_{14} + G(0, 0.015)$
	D_2O & Salt	All	$\beta_{14}' = (1 + 0.00276\ (T_{eff} - T_{eff}^N))\ \beta_{14}$

Table 8.7: Summary of the corrections required for the observable parameters. $G(0,X)$ represents a smearing of the distribution by a Gaussian of width X. Values taken from [47]. T_{eff}^N is the energy of the ^{16}N calibration source. Explicit functional forms for the energy corrections can be found in [51] and [47].

8.2.2 Flux Corrections

As discussed in section 7.1.3, the predicted neutrino event rate in the SNO detector was used to constrain the numbers of neutrino interactions occurring in the two phases in order to fit for a single interaction rate for each signal type across the two phases. The Monte Carlo simulation was used to determine the expected number of CC, ES and NC events according to the SSM, and the interaction rates to be fit were formulated in terms of the observed fraction of this flux. Several corrections had to be applied in order for the Monte Carlo prediction to best model reality. These are described below.

The predicted number of events for signal type i, accounting for all correction factors, is given by:

$$N_i = N_i^{MC} \, \mathcal{F}^{ssm} \, \delta^{sim} \, \delta_i^{sac} \, N_i^{iso} \, N_i^D \, N_i^e \, \tau \tag{8.14}$$

where:

- N_i^{MC} is the number of events predicted by the Monte Carlo simulation for signal i for an incident flux of one SSM, assuming an unoscillated, undistorted incident ν_e flux. This is recalculated for each signal extraction and therefore accounts for any systematic shifts applied to the PDFs.

- $\mathcal{F}^{ssm}$ is a correction to the SSM rate used by the simulation. The Monte Carlo events were generated using the BP2000 model flux of $5.15 \times 10^6 \, \mathrm{cm}^{-2} \, \mathrm{s}^{-1}$ [81]. This factor converts to the BS2005 model flux of $5.69 \times 10^6 \, \mathrm{cm}^{-2} \, \mathrm{s}^{-1}$ [13].

- δ^{sim} accounts for events aborted in the simulation due to photon tracking errors. This correction was evaluated differentially with energy since it increases with the number of photons in an event. This correction is derived in appendix E.

- δ_i^{sac} corrects for differences in the acceptance of the instrumental and high level cuts for data and Monte Carlo events. Full details of this can be found in chapter 6.

- N_i^{iso} is a correction to account for CC interactions on chlorine and sodium nuclei in the D_2O volume, which are not modelled in the Monte Carlo simulation. This correction is only relevant to the CC signal.

- N_i^D is a correction to the number of target deuterons and hence is relevant to CC and NC only.

- N_i^e is a correction to the number of target electrons and hence is relevant to ES only.

- τ corrects for deadtime introduced into the data set by the instrumental cuts, as discussed in section 6.4.

The values for the corrections are summarised in table 8.8.

Correction	CC	ES	NC
SSM flux	1.105	1.105	1.105
Aborted MC events	see text	see text	see text
Sacrifice (D_2O)	0.9924	0.9924	0.9924
Sacrifice (Salt)	0.9930	0.9930	0.9954
Cl & Na isotopes	1.0002	—	—
Target deuterons	1.0122	—	1.0122
Target electrons	—	1.0131	—
Livetime (D_2O)	0.986	0.986	0.986
Livetime (Salt)	0.989	0.989	0.989

Table 8.8: Summary of the corrections applied to the expected number of CC, ES and NC events used in the signal extraction.

8.2.3   *hep* Flux Correction

The SNO data set also contains some contribution from *hep* neutrinos. The flux of *hep* events within the energy range of interest is very low in comparison to the ^{8}B flux (see table 1.1), such that the exact value has a negligible impact on the results. Hence, for simplicity, the Monte Carlo simulation was used to fix the expected numbers of these events. PDFs were created for these events and incorporated into the signal extraction in order to model the shapes of the distributions of these events accurately, so that they were appropriately "subtracted" from the CC, ES and NC flux spectra.

Additional backgrounds to the signal were expected from atmospheric neutrino events and from radioactive background events originating from contamination on the acrylic vessel [47]. Atmospheric neutrino events typically have energies measured in GeV and, so, background events occurring within the energy range of interest are very rare. The contribution of these event types to the data set is known to be very low, on the order of a few events, and so for the purposes of this thesis, these were taken to be negligible. The Leslie events have characteristics very similar to the external background events (specifically those originating on the AV or in the H_2O region) and so were assumed to fit out as those event types in the extraction. Any correction to the neutrino fluxes was taken to be negligible.

8.3 Neutrino Flux and Spectrum Results

This section describes the results of the signal extraction process performed on the full LETA data set, including the full propagation of systematic uncertainties. The analysis presented here was performed using the FTK energy fitter and the FTP reconstruction algorithm (section 3.3.2).

The neutrino fluxes were extracted both assuming an undistorted ^{8}B neutrino spectrum (a 'constrained' fit) and a fit in which the shape of the reconstructed spectra were allowed to vary ('unconstrained'). In the constrained fits, the CC and ES spectra were constrained to that predicted from an undistorted spectrum and the normalisation was treated as a single fit parameter. In the unconstrained fits, the apparent number of CC and ES events in each energy bin were treated as independent parameters, thus allowing the shape of the spectra to vary. The full propagation of systematic uncertainties was performed in both fit types and the results are presented here.

As discussed in section 8.1, a new minimum was found when scanning the likelihood space in the systematic parameters and, so, the PDFs used in the signal extraction were corrected to account for this. The final signal extraction was performed at this point. The corrections applied to the PDFs were a shift of +0.75% in the energy scale, a shift of 0.3% in the electron β_{14} scale and an additional energy smearing of 0.02 MeV and 0.025 MeV in the D$_2$O and salt phases, respectively. The results presented below are for the extraction performed at this point, using the values of systematic uncertainties extracted from the likelihood scans, as given in section 8.1 and appendix F.

8.3.1 Neutrino Fluxes

Implementing the corrections given in section 8.2 and applying the systematic uncertainties as discussed in section 8.1, the full signal extraction was performed in both an energy-constrained and an unconstrained fit. Although the constrained fit assumed an undistorted spectrum, any predicted distortion (such as that from the MSW effect, as discussed in chapter 1) is expected to be very small (on the order of 8%) and, so, this is a reasonable approximation for the purposes of extracting the integral flux values. The CC and ES fluxes are quoted for the constrained fit only, since the evaluation of an integral flux involves an assumption about the shape of the spectrum outside the energy range under consideration. This assumption is implicit in the constrained fit, whereas no information is available with which to extrapolate the shape of the unconstrained spectra. The NC flux extracted from the two types of fit can be compared to each other since the shape of the neutron energy spectrum is well understood. The fluxes extracted using a 3.5 MeV threshold in T_{eff} were:

Constrained Fit

$$\Phi_{\text{CC}}^{con} = 1.622\,^{+0.038}_{-0.038}(\text{stat})\,^{+0.034}_{-0.034}(\text{syst}) \times 10^6\,\text{cm}^{-2}\,\text{s}^{-1} \qquad (8.15)$$

$$\Phi_{\text{ES}}^{con} = 2.444\,^{+0.163}_{-0.163}(\text{stat})\,^{+0.068}_{-0.072}(\text{syst}) \times 10^6\,\text{cm}^{-2}\,\text{s}^{-1} \qquad (8.16)$$

$$\Phi_{\text{NC}}^{con} = 5.194\,^{+0.139}_{-0.140}(\text{stat})\,^{+0.106}_{-0.102}(\text{syst}) \times 10^6\,\text{cm}^{-2}\,\text{s}^{-1} \qquad (8.17)$$

Unconstrained Fit

$$\Phi_{\text{NC}}^{uncon} = 5.191\,^{+0.156}_{-0.155}(\text{stat})\,^{+0.132}_{-0.099}(\text{syst}) \times 10^6\,\text{cm}^{-2}\,\text{s}^{-1} \qquad (8.18)$$

These results are in very good agreement with each other. In addition, the NC flux shows good agreement with the solar neutrino flux predicted by the SSM of $5.69 \pm 0.85 \times 10^6\,\text{cm}^{-2}\,\text{s}^{-1}$ [13]. Both the CC and ES fluxes are significantly suppressed relative to the total predicted solar neutrino flux, demonstrating that the electron neutrinos produced in the Sun's core are oscillating to other flavours en route to the Earth.

The results from previous analyses of SNO data are summarised in table 8.9. The results presented above are in good agreement with those from previous analyses, but with significantly reduced uncertainties. The improvements made for this analysis are particularly evident in the systematic uncertainties, which have been significantly reduced in comparison to previous results. This is a result of the improvements made to the detector model and to the analysis of systematic uncertainties performed prior to the signal extraction (such as those presented in chapters 5 and 6) and is also due to the more accurate handling of these uncertainties, utilising the likelihood values in the extraction process. The statistical uncertainties have also been reduced in comparison to previous results due to the combination of two phases of data and the lowered energy threshold.

	CC	ES	NC
D$_2$O constrained	$1.76\,^{+0.06\ +0.09}_{-0.05\ -0.09}$	$2.39\,^{+0.24\ +0.12}_{-0.23\ -0.12}$	$5.09\,^{+0.44\ +0.46}_{-0.43\ -0.43}$
Salt constrained	$1.72\,^{+0.05\ +0.11}_{-0.05\ -0.11}$	$2.34\,^{+0.23\ +0.15}_{-0.23\ -0.14}$	$4.81\,^{+0.19\ +0.28}_{-0.19\ -0.27}$
Salt unconstrained	$1.68\,^{+0.06\ +0.08}_{-0.06\ -0.09}$	$2.35\,^{+0.22\ +0.15}_{-0.22\ -0.15}$	$4.94\,^{+0.21\ +0.38}_{-0.21\ -0.34}$

Table 8.9: Summary of flux measurements from previous SNO publications [12]. All units are $\times 10^6\,\text{cm}^{-2}\,\text{s}^{-1}$. The quoted uncertainties are statistical and systematic respectively.

In principle, the ratio of the CC to NC fluxes can be used as a measure of the observed ν_e suppression, independent of any theoretical prediction. The correlation between the parameters can be extracted directly from the fit and this was assumed to describe the correlations for both statistical and systematic uncertainties. Although some of the systematic uncertainties would cancel in the ratio (such as the position uncertainties, which

could be assumed to have the same effect for both interaction types) the contribution of these effects to the overall uncertainty is negligible and, so, the full systematic uncertainties were used in this calculation. Using the constrained fit results for both fluxes for consistency, and the extracted correlation of -0.3536, the resulting ratio is:

$$\frac{\Phi_{\text{CC}}^{con}}{\Phi_{\text{NC}}^{con}} \; = \; 0.312\,^{+0.013}_{-0.013}(\text{stat})\,^{+0.011}_{-0.010}(\text{syst}) \tag{8.19}$$

which is a strong indication of the precision with which the suppression of the ν_e survival probability due to MSW effects in the Sun has been measured. It is also additional evidence for this effect since vacuum oscillation alone results in a maximum suppression of a factor of two. This value is in good agreement with the prediction for the LMA MSW scenario and with previously published SNO results (equation 1.20) but with significantly reduced systematic and statistical uncertainties.

The one-dimensional projections of the observable distributions are shown in figures 8.1 and 8.2 for the unconstrained fit. The total fit result is the sum of the individual Monte Carlo PDFs, normalised to the number of events obtained from the signal extraction for each one. The χ^2 values for the fits are given in table 8.10, for both the constrained and unconstrained fits. The χ^2 values are in fact upper bounds on the goodness of fit, since they were evaluated using only the statistical uncertainties on the data, not accounting for systematic or statistical uncertainties on the PDFs and, hence, have underestimated the total uncertainty. The χ^2 values show that the results from the signal extraction give good fits to the data for all four observable distributions in each of the two phases, for both fit types. In addition, the χ^2 values are quoted for the fits at the original point, before the application of the shifts derived using the likelihood scans in section 8.1. Although the caveats mentioned above limit the interpretation of the absolute values, the improvement in the β_{14} fit in particular can be seen from a comparison of the two sets of χ^2 results.

The number of background events extracted from the fits were converted into rates using the simulation, to account for the difference in acceptance between the two phases due to livetime and degrading detector resolution. The results are given in table 8.11. Previous analyses observed a significant difference in the extracted rates between the constrained and unconstrained fits, in particular for internal ^{214}Bi events. This was evidence of a bias that had been observed in the unconstrained fit. The good agreement shown here between the two fits suggests that the spectra extracted from the unconstrained fit are consistent at the one sigma level with the undistorted spectra assumed in the constrained fit and that the level of bias present in the fit is minimal, as claimed in chapter 7. A comparison of the extracted rates between the two phases highlights a few discrepancies, particularly in the external backgrounds. This is partly explained by the *ex-situ* con-
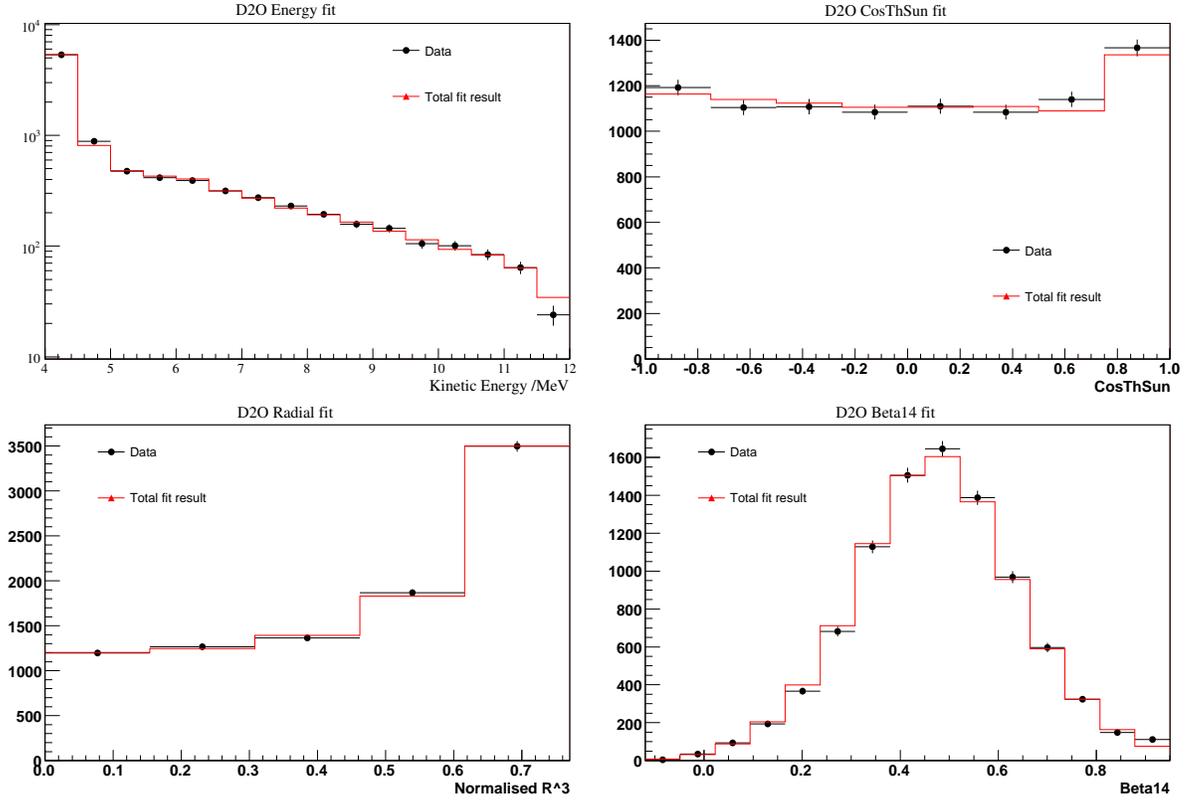
Figure 8.1: Fit results to the observable distributions for the D_2O phase data set, in a joint-phase unconstrained fit.

straints (table 3.2), which show that the level of contamination in the H_2O was lower in the salt phase. The large difference in the PMT rates is misleading since the knowledge of the energy scale near the PMTs is very poor. This leads to a large uncertainty in the conversion to rate, on the order of 50% [82], due to the steepness of the PMT energy spectrum. It is also known that there is some residual bias in distinguishing different backgrounds in the fit, which could differ between the phases. However, the ensemble tests presented in chapter 7 demonstrated that this does not affect the neutrino fluxes.

The dominant systematic uncertainties on the CC and NC fluxes in the constrained fit are the energy non-linearity, the energy-dependent fiducial volume and the sacrifice measurements and neutron capture in the salt phase for the NC flux. Previous analyses saw much greater effects from the energy scale, energy resolution and β_{14} uncertainties, but these have been reduced in this analysis by the use of the likelihood to extract more accurate values for these uncertainties. The uncertainty on the ES flux is dominated by the angular resolution. The individual contributions from each uncertainty to the total systematic uncertainty on the neutrino fluxes are listed in tables 8.12 – 8.14.
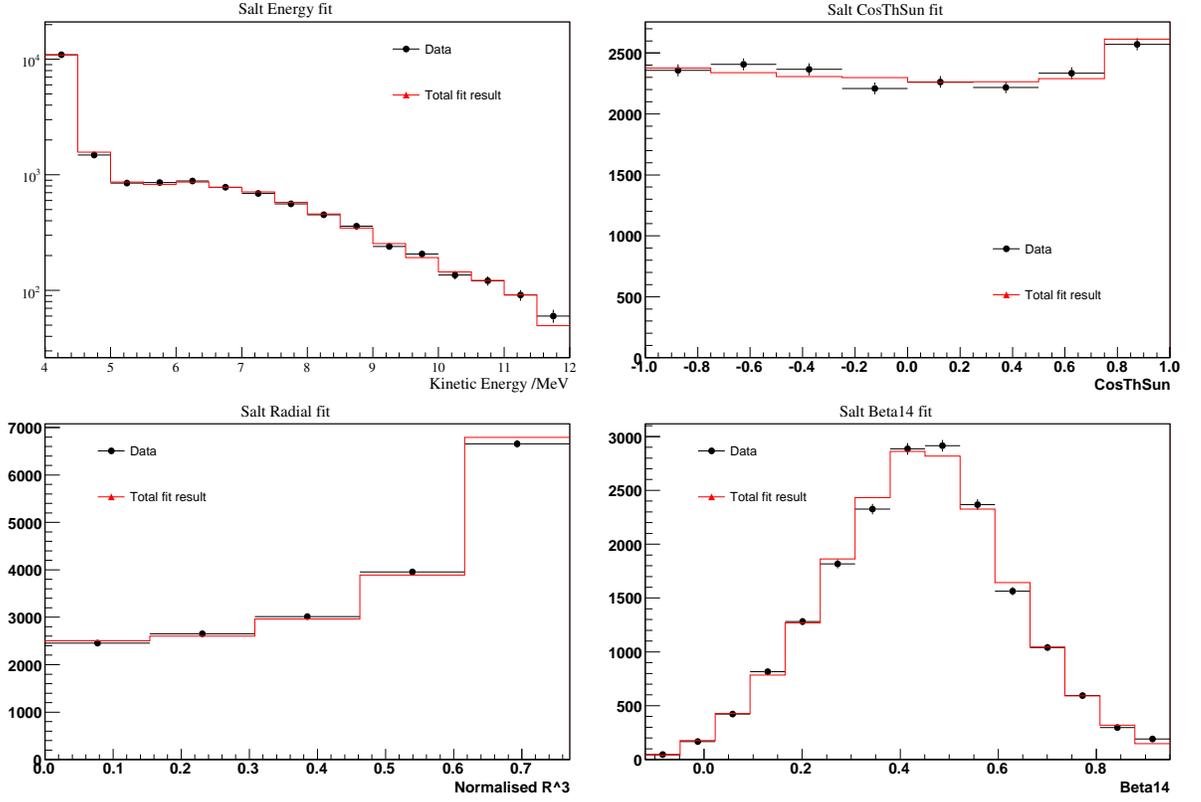
Figure 8.2: Fit results to the observable distributions for the salt phase data set, in a joint-phase unconstrained fit.

Phase	Observable	χ^2/dof			
		Unconstrained		Constrained	
		Final	Original	Final	Original
D$_2$O	T_{eff}	0.957 (15)	1.070 (15)	1.283 (17)	1.373 (17)
	$\cos\theta_\odot$	0.860 (7)	0.871 (7)	0.994 (7)	1.007 (7)
	ρ	0.405 (4)	0.359 (4)	0.372 (4)	0.336 (4)
	β_{14}	1.663 (14)	1.844 (14)	1.660 (14)	1.812 (14)
Salt	T_{eff}	0.831 (15)	0.795 (15)	1.435 (17)	1.238 (17)
	$\cos\theta_\odot$	1.387 (7)	1.382 (7)	1.390 (7)	1.392 (7)
	ρ	1.629 (4)	1.701 (4)	1.590 (4)	1.700 (4)
	β_{14}	1.956 (14)	1.995 (14)	2.022 (14)	2.048 (14)

Table 8.10: χ^2 values for the fit of the extracted signals to the data set for the four observable distributions used in the signal extraction, in each of the two phases. 'Final' refers to the final signal extraction, including the shifts applied to the PDFs in energy and β_{14}. 'Original' refers to the signal extraction performed before these shifts were applied. The χ^2 values are given for both the constrained and unconstrained fits, quoted per degree of freedom, with the number of degrees of freedom given afterwards in parentheses.

Phase	Background type	Constrained	Unconstrained
D_2O	AV neutrons /mHz	$0.145 \pm 0.044^{+0.026}_{-0.019}$	$0.141 \pm 0.049^{+0.041}_{-0.017}$
	D_2O ^{214}Bi/mHz	$95.8 \pm 5.93^{+14.7}_{-8.32}$	$99.0 \pm 8.83^{+19.7}_{-12.3}$
	D_2O ^{208}Tl/mHz	$0.530 \pm 0.277^{+0.338}_{-0.306}$	$0.527 \pm 0.274^{+0.321}_{-0.298}$
	AV ^{214}Bi/Hz	$3.42 \pm 0.487^{+0.423}_{-0.883}$	$3.511 \pm 0.499^{+0.417}_{-0.907}$
	AV ^{208}Tl/mHz	$26.2 \pm 16.9^{+21.6}_{-24.8}$	$27.7 \pm 17.8^{+22.9}_{-25.1}$
	H_2O ^{214}Bi/Hz	$9.28 \pm 1.42^{+2.00}_{-0.461}$	$9.25 \pm 1.42^{+2.02}_{-1.63}$
	H_2O ^{208}Tl/Hz	$0.306 \pm 0.083^{+0.060}_{-0.043}$	$0.305 \pm 0.084^{+0.062}_{-0.061}$
	PMT β-γs /kHz	$6.87 \pm 0.330^{+0.498}_{-0.377}$	$6.75 \pm 0.344^{+0.566}_{-0.414}$
Salt	AV neutrons /mHz	$0.000 \pm 0.000^{+0.001}_{-0.000}$	$0.000 \pm 0.000^{+0.004}_{-0.000}$
	D_2O ^{214}Bi/mHz	$126.0 \pm 6.02^{+10.9}_{-9.70}$	$129.0 \pm 8.54^{+15.7}_{-14.0}$
	D_2O ^{208}Tl/mHz	$1.76 \pm 0.501^{+0.425}_{-0.391}$	$1.704 \pm 0.505^{+0.429}_{-0.422}$
	AV ^{214}Bi/Hz	$0.654 \pm 0.354^{+0.426}_{-0.366}$	$0.726 \pm 0.350^{+0.472}_{-0.347}$
	AV ^{208}Tl/mHz	$65.5 \pm 12.0^{+8.02}_{-7.66}$	$65.6 \pm 12.0^{+8.06}_{-9.50}$
	H_2O ^{214}Bi/Hz	$6.39 \pm 1.40^{+0.460}_{-0.450}$	$6.40 \pm 1.40^{+0.800}_{-1.25}$
	H_2O ^{208}Tl/Hz	$0.166 \pm 0.050^{+0.013}_{-0.008}$	$0.167 \pm 0.050^{+0.022}_{-0.020}$
	PMT β-γs /kHz	$2.89 \pm 0.082^{+0.116}_{-0.158}$	$2.87 \pm 0.084^{+0.125}_{-0.169}$
	^{24}Na /mHz	$0.231 \pm 0.075^{+0.016}_{-0.019}$	$0.236 \pm 0.075^{+0.015}_{-0.026}$

Table 8.11: Background event rates obtained from the signal extraction in each phase. Uncertainties quoted are statistical and systematic, respectively.

Systematic uncertainty /%	CC	ES	NC
D_2O Angular resolution (+)	-0.164	1.48	-0.038
D_2O Angular resolution (-)	0.173	-1.57	0.042
Salt Angular resolution (+)	-0.21	1.87	-0.094
Salt Angular resolution (-)	0.223	-2.03	0.105
β_{14} electron scale (+)	-0.232	-0.181	0.161
β_{14} electron scale (-)	0.12	0.137	-0.059
β_{14} neutron scale (+)	-0.33	-0.24	0.435
β_{14} neutron scale (-)	0.204	0.14	-0.271
β_{14} electron width (+)	-0.102	-0.029	0.102
β_{14} electron width (-)	0.151	0.0667	-0.168
β_{14} neutron width (+)	-0.032	-0.028	0.046
β_{14} neutron width (-)	0.036	0.028	-0.060
β_{14} energy dep (+)	-0.132	-0.004	0.241
β_{14} energy dep (-)	0.121	0.054	-0.221

Table 8.12: Contributions of individual systematic uncertainties on β_{14} and $\cos\theta_\odot$ to the total uncertainty on the CC, ES and NC fluxes extracted from the constrained fit. All values are quoted in %.

Systematic uncertainty /%	CC	ES	NC
Energy scale (cor) (+)	-0.274	-0.080	0.136
Energy scale (cor) (-)	0.314	0.143	-0.184
D_2O Energy scale (uncor) (+)	-0.442	-0.348	0.251
D_2O Energy scale (uncor) (-)	0.079	0.122	-0.038
Salt Energy scale (uncor) (+)	-0.079	0.012	0.053
Salt Energy scale (uncor) (-)	0.565	0.0645	-0.456
Energy scale (neut) (+)	-0.083	0.093	-0.102
Energy scale (neut) (-)	0.130	-0.096	0.104
D_2O Energy resolution (+)	-0.036	-0.268	0.053
D_2O Energy resolution (-)	-0.123	0.322	0.233
Salt Energy resolution (+)	-0.001	0.002	0.077
Salt Energy resolution (-)	0.086	0.160	-0.468
Energy non-linearity	0.609	-0.129	-0.580
D_2O X scale (+)	0.034	0.028	-0.020
D_2O X scale (-)	-0.277	-0.383	0.066
Salt X scale (+)	0.017	0.015	0.039
Salt X scale (-)	-0.16	-0.154	-0.15
D_2O Y scale (+)	0.037	0.041	-0.014
D_2O Y scale (-)	-0.258	-0.282	0.079
Salt Y scale (+)	0.019	0.031	0.015
Salt Y scale (-)	-0.11	-0.145	-0.105
D_2O Z scale (+)	0.236	0.225	-0.056
D_2O Z scale (-)	-0.214	-0.241	0.077
Salt Z scale (+)	0.024	0.033	0.044
Salt Z scale (-)	-0.26	-0.268	-0.383
D_2O X offset (+)	0.003	-0.010	-0.002
D_2O X offset (-)	-0.006	-0.002	0.001
Salt X offset (+)	-0.019	-0.009	-0.026
Salt X offset (-)	-0.004	0.008	0.007
D_2O Y offset (+)	0.011	-0.048	-0.014
D_2O Y offset (-)	-0.006	-0.017	-0.004
Salt Y offset (+)	0.011	-0.036	-0.040
Salt Y offset (-)	-0.004	0.008	0.016
D_2O Z offset (+)	-0.016	-0.088	-0.003
D_2O Z offset (-)	-0.003	0.013	-0.004
Salt Z offset (+)	-0.001	-0.021	-0.119
Salt Z offset (-)	0.004	0.031	-0.006
D_2O X resolution	-0.030	-0.092	-0.024
Salt X resolution	-0.001	0.001	-0.066
D_2O Y resolution	-0.007	0.043	-0.014
Salt Y resolution	-0.020	-0.033	-0.088
D_2O Z resolution	-0.005	0.025	-0.010
Salt Z resolution	0.007	-0.072	0.135

Table 8.13: Contributions of individual systematic uncertainties on T_{eff} and ρ to the total uncertainty on the CC, ES and NC fluxes extracted from the constrained fit. All values are quoted in %.

Systematic uncertainty /%	CC	ES	NC
^{24}Na model (+)	0.005	-0.009	0.014
^{24}Na model (-)	0.008	-0.020	0.079
D_2O Energy dependent fiducial volume (+)	-1.56	-0.733	0.707
D_2O Energy dependent fiducial volume (-)	0.918	0.424	-0.412
Salt Energy dependent fiducial volume (+)	-0.875	-0.365	-0.362
Salt Energy dependent fiducial volume (-)	1.04	0.428	0.422
Sacrifice (+)	-0.474	-0.48	-0.387
Sacrifice (-)	1.14	1.13	1.04
Photodisintegration (+)	-0.001	0.000	-0.168
Photodisintegration (-)	0.001	0.000	0.170
D_2O Neutron capture (+)	-0.023	-0.023	-0.066
D_2O Neutron capture (-)	0.024	0.024	0.065
Salt Neutron capture (+)	0.023	0.023	-1.22
Salt Neutron capture (-)	-0.022	-0.023	1.25
D_2O PMT energy slope (+)	-0.024	0.036	0.042
D_2O PMT energy slope (-)	-0.001	-0.000	-0.023
Salt PMT energy slope (+)	0.026	-0.033	0.341
Salt PMT energy slope (-)	-0.067	-0.062	-0.568
D_2O PMT radial slope (+)	0.011	-0.081	-0.090
D_2O PMT radial slope (-)	-0.027	0.066	0.111
Salt PMT radial slope (+)	-0.028	-0.279	-0.181
Salt PMT radial slope (-)	0.023	0.342	0.120
D_2O PMT radial offset (+)	-0.027	0.054	0.087
D_2O PMT radial offset (-)	0.074	-0.275	-0.225
Salt PMT radial offset (+)	0.021	0.252	0.167
Salt PMT radial offset (-)	-0.044	-0.385	-0.326
D_2O PMT β_{14} mean (+)	0.036	0.021	-0.045
D_2O PMT β_{14} mean (-)	-0.042	-0.065	0.047
Salt PMT β_{14} mean (+)	0.104	0.079	-0.123
Salt PMT β_{14} mean (-)	-0.101	-0.101	0.102
D_2O PMT β_{14} width (+)	-0.017	-0.069	0.022
D_2O PMT β_{14} width (-)	0.023	0.027	-0.022
Salt PMT β_{14} width (+)	-0.055	-0.085	0.305
Salt PMT β_{14} width (-)	0.057	0.067	-0.344

Table 8.14: Contributions of individual systematic uncertainties on the PDF shapes to the total uncertainty on the CC, ES and NC fluxes extracted from the constrained fit. All values are quoted in %.

8.3.2   CC and ES Energy Spectra

The CC and ES energy spectra extracted from the unconstrained fit are shown in figures 8.3 and 8.5, in terms of the observed fraction of the unoscillated SSM. An undistorted energy spectrum with no oscillation or suppression would therefore be flat and of magnitude 1.0. The statistical uncertainties are represented by the error bars and the systematic uncertainties on the flux in each energy bin are represented by the coloured bands. The constrained fit result is superimposed on each figure as a dashed line, for comparison.

The uncertainties can be seen to increase as the energy threshold is lowered. This is due to the large number of background events, which dominate the measurement at these low energies. However, the inclusion of the low energy bins serves to normalise the background levels, reducing the uncertainties in bins further up the spectrum and, thus, allowing a more accurate extraction of the spectral shape.

For comparison, the CC spectrum resulting from a previous analysis of the salt phase data down to $5.5 \, \text{MeV}$ is shown in figure 8.4. The improvement in the evaluation of the systematic uncertainties for LETA can be seen from a comparison of the coloured bands on these figures. Since this is still a preliminary analysis, the uncertainties can be expected to reduce further for the final results. In particular, verification of the true minimum in the correlated energy scale and, thus, a reduction in the associated uncertainty will have a big impact on the uncertainties in the shape of the CC spectrum. Further investigation into the true shape of the PMT PDF and tests of the validity of floating some aspects of the PDF shape in the fit could also have a noticeable effect on the final uncertainties.

The effects of the individual systematic uncertainties on the shape of the CC and ES spectra are shown in figures 8.6 and 8.7, respectively. For some systematic parameters, both the positive and the negative value of the uncertainty caused the extracted neutrino flux to be shifted in the same direction. For these parameters, a conservative approach was taken in which the maximum of the two shifts was used as the systematic uncertainty due to that parameter. The dominant uncertainties on the CC spectrum are the energy scale and resolution and the shape of the analytic form used for the PMT PDF. For the ES spectrum, the angular resolution is a more dominant uncertainty since the shape of the $\cos \theta_\odot$ PDF for ES was one of the strongest tools for separation. However, the low statistics of the flux dominate the ES uncertainty.

The spectrum shown in figure 8.5 seems to suggest a systematic rise at low energies. As discussed in section 7.1.1, the unconstrained fit used PDFs that were created assuming an undistorted spectrum. Therefore, the uncertainties on the resulting spectra cannot be directly utilised in a quantitative test of the MSW prediction. However, a χ^2 fit can be performed to the prediction for an undistorted spectrum which, as discussed, would be a flat line on figure 8.5. Including the systematic uncertainties in this calculation would

Figure 8.3: The extracted spectrum for CC events shown in terms of the fraction of the SSM predicted flux. The errors bars represent the statistical fit uncertainties and the coloured band represents systematic uncertainties.

Figure 8.4: CC spectrum from a previous analysis of the full salt phase data set down to 5.5 MeV, shown in terms of the fraction of the SSM prediction. The errors bars represent statistical uncertainties and the coloured band represents systematic uncertainties.
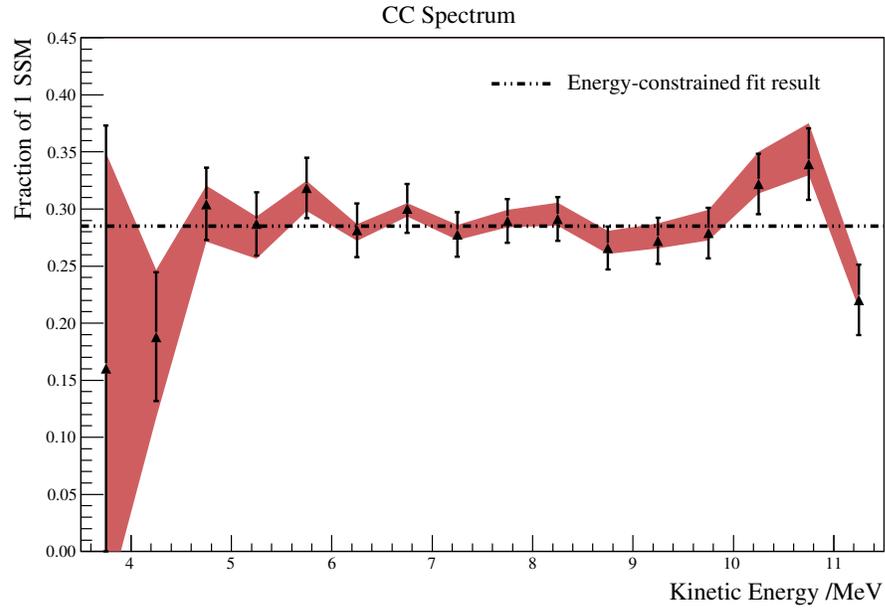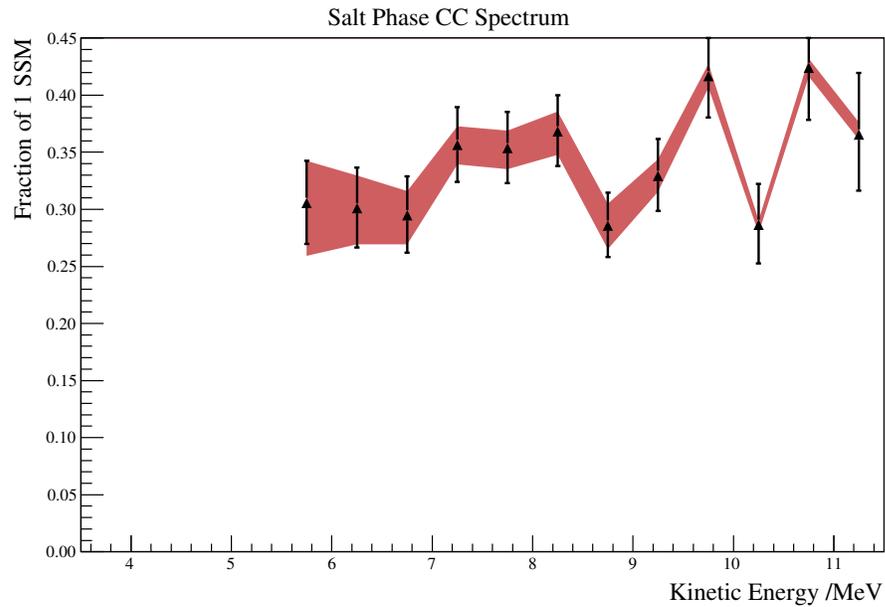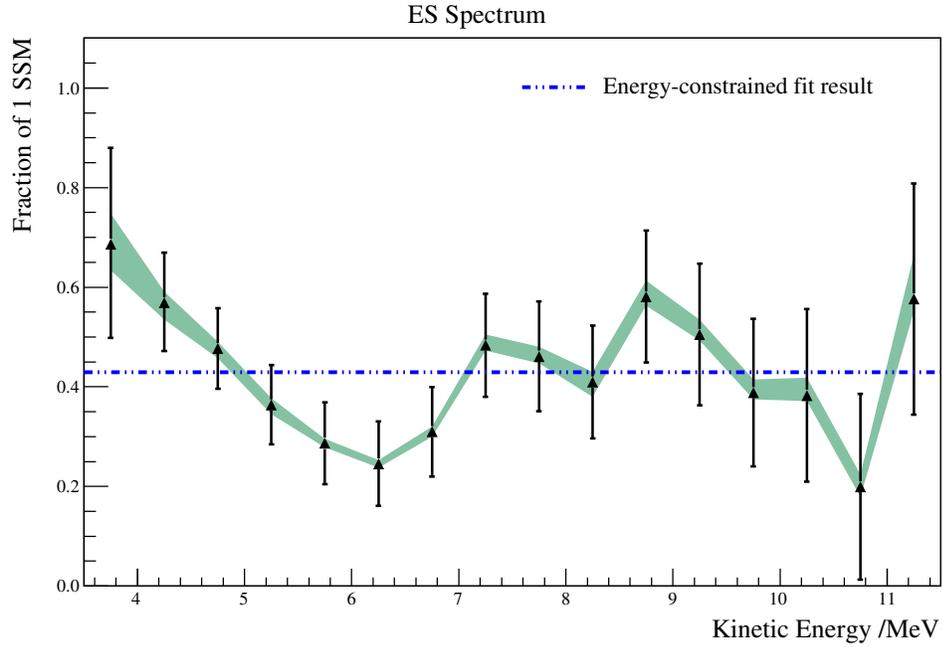
Figure 8.5: The extracted spectrum for ES events shown in terms of the fraction of the SSM predicted flux. The errors bars represent the statistical fit uncertainties and the coloured band represents systematic uncertainties.

improve the goodness-of-fit. However, only the systematics that have a differential effect on the spectrum would apply, not those that only affect the normalisation and, so, for the purposes of this calculation only the statistical uncertainties were used. Taking the integral ES flux from the constrained fit to define the prediction for the undistorted case, the evaluated χ^2 for the fit was found to be 18.2 over 15 degrees of freedom. This has a confidence level of 25%. Therefore, the rise is not strongly significant.

In order to quantitatively test predictions for possible distortions in the spectrum, a method was devised to explore the possible values of the MSW parameters and to find the best-fit point and associated uncertainties utilising the likelihood value from the signal extraction. This is described in the following section.

8.4 Hypothesis Testing

The predicted suppression of the electron neutrino survival probability can be used to test the model of MSW-enhanced oscillation for particular values of the oscillation parameters. The SNO Monte Carlo was generated assuming an undistorted neutrino spectrum, with no oscillation. For a particular point in the $[\theta_{12}, \Delta m_{12}^2]$ parameter space, the ν_e survival probability can be calculated as a function of neutrino energy. This can then be applied
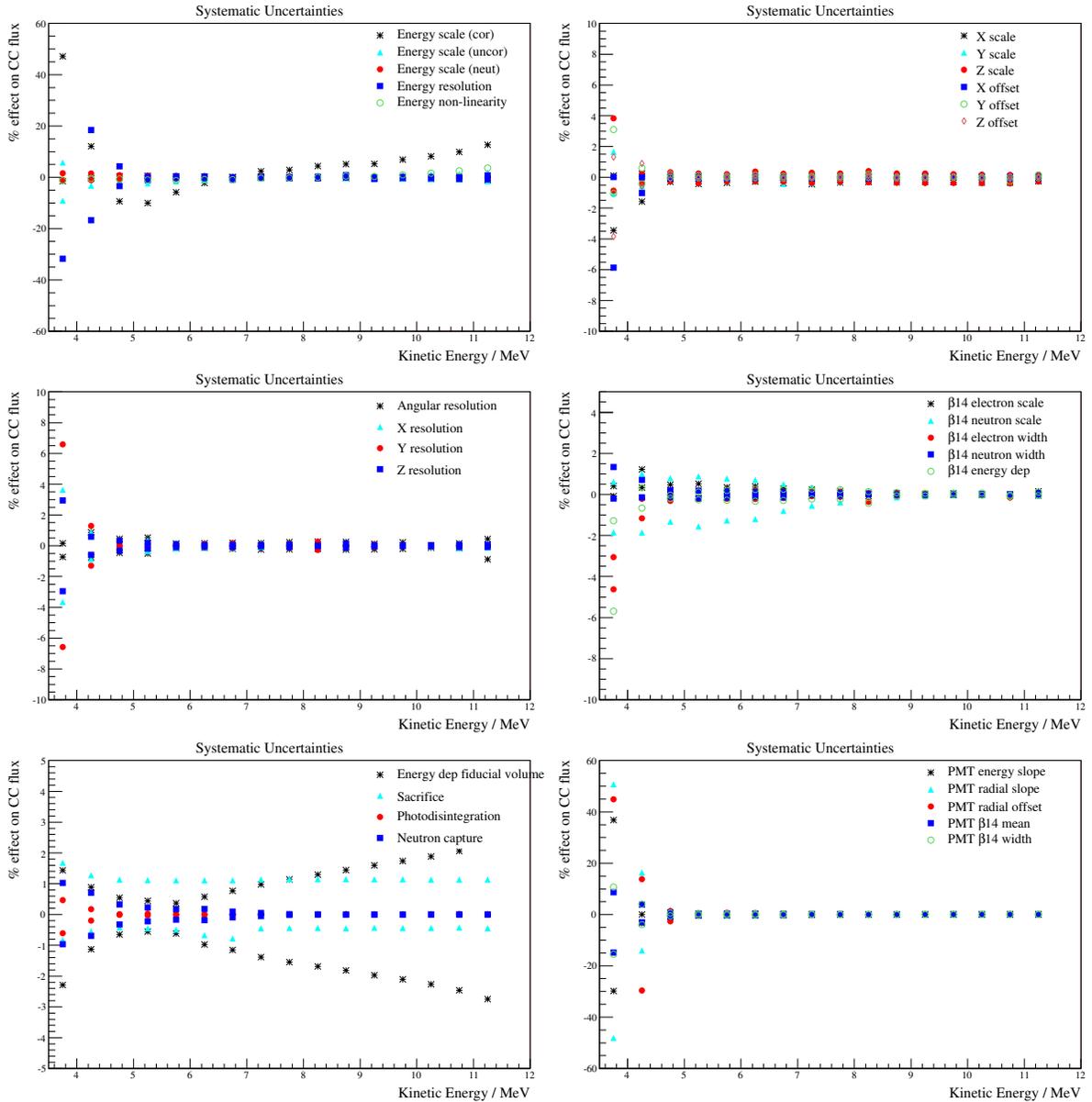
Figure 8.6: The effect of individual systematic uncertainties on the CC energy spectrum. Where uncertainties were treated as uncorrelated between the phases, the effect of the uncertainty in the two phases has been summed in quadrature.

to the Monte Carlo distributions as a suppression to the CC and ES spectra in order to model the MSW prediction at that point. For the CC spectra, the survival probability was applied as a direct weighting to every event since the CC interaction in SNO is only sensitive to ν_e. For ES events, the weighting takes into account the limited sensitivity to other flavours, as discussed in chapter 1.
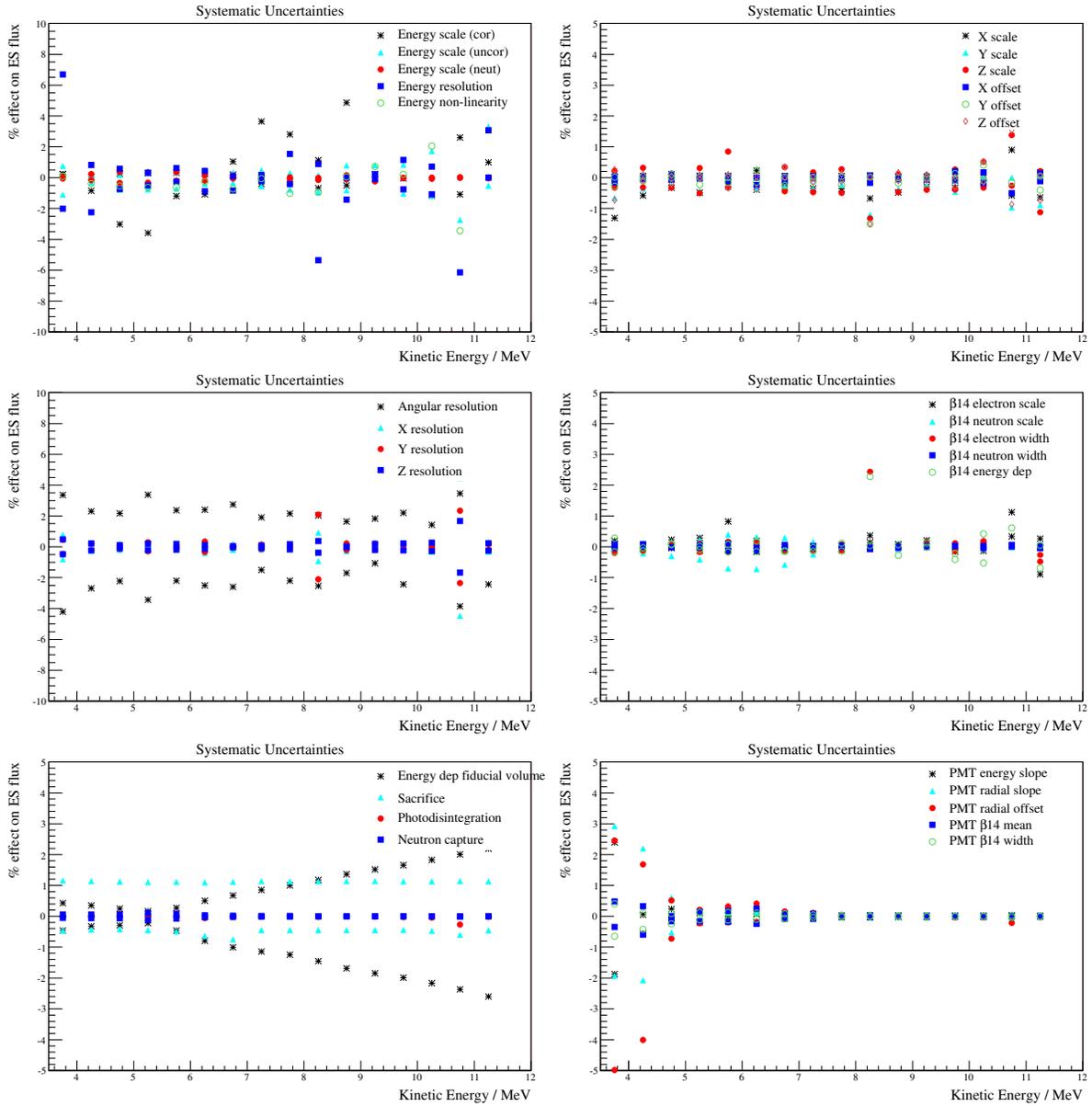
Figure 8.7: The effect of individual systematic uncertainties on the ES energy spectrum. Where uncertainties were treated as uncorrelated between the phases, the effect of the uncertainty in the two phases has been summed in quadrature.

An energy-constrained fit can then be performed at each point on the MSW plane, allowing just the total neutrino flux to vary to account for the theoretical uncertainty in the ^{8}B flux. Both the shapes of the individual neutrino spectra and the relative normalisation are fixed by the theoretical prediction at that point.

The likelihood for each fit can be used to find the best-fit point in parameter space

and also to evaluate the contours that represent the uncertainties in that best-fit point. As discussed in section 7.5.2, the one sigma uncertainty is given by the points at which the log likelihood has decreased by 0.5 from the global maximum. Extending the argument, it can be seen from equation 7.21 that the contour representing the nth sigma bound on the parameters is given by the points at which:

$$\mathbb{L}(\mu_i) - \mathbb{L}_{max}(\mu_i) = -\frac{n^2}{2} \tag{8.20}$$

i.e. the log likelihood has decreased by $\frac{n^2}{2}$ from the global maximum.

Previous analyses used a χ^2 method to evaluate the contours. This involved using the survival probability to calculate the neutrino spectrum at each point in the parameter space and then convolving that with the the cross sections for each neutrino interaction to get the resulting recoil electron spectra. This was then further convolved with detector resolution functions to predict the shapes of the reconstructed energy spectra. A χ^2 fit was then performed of those predicted spectra to the CC and ES spectra obtained from the signal extraction and the resulting χ^2 values for each point in the MSW plane were used to find the best-fit point and to evaluate the contours.

The use of the unconstrained spectrum from signal extraction involved some degree of model assumption, since the PDFs used in the extraction were generated assuming no oscillations and an undistorted spectrum. In addition, the uncertainties on that spectrum are not straightforward to interpret, particularly at points in the plane where a more significant distortion would be predicted. The use of the likelihood function as described above is an improvement over this method since the PDFs were generated according to the predicted suppression at each point. In addition, the full information about all four observable parameters is used in the fit at each point, whereas the χ^2 method uses only the energy spectra.

8.4.1 Scanning the MSW Plane

Using the method described above, a scan was performed of the region of the MSW plane containing the current preferred values for the solar neutrino mixing parameters corresponding to the LMA solution, as discussed in chapter 1. The values of the ν_e survival probability were provided by Simard [83]. The region scanned was [-8.0, -3.0] in $\log \Delta m_{12}^2$, in steps of size 0.05 each, and [0.05, 0.95] in $\tan^2 \theta_{12}$ in steps of 0.1. A finer gridded scan was performed around the best-fit point, sampling both $\log \Delta m_{12}^2$ and $\tan^2 \theta_{12}$ in steps of 0.02.

The result of this scan is shown in figure 8.8. The first four contours are shown and the best-fit point is marked with a star. The best-fit point was found to be at:

$$\Delta m_{12}^2 = 6.03 \times 10^{-5} \, \text{eV}^2$$
$$\tan^2 \theta_{12} = 0.42$$

This result is consistent to within $2\,\sigma$ with the LMA best-fit point from a previous global analysis of solar data given in section 1.2.4 [12]. In addition, the value of the likelihood extracted for this point shows that it is consistent at a level of approximately $2\,\sigma$ with the result from the constrained fit described in section 8.3, which assumed an undistorted spectrum. This demonstrates that the data is consistent both with the predicted MSW-rise (as given by the LMA best-fit point) and an undistorted spectrum at a $2\,\sigma$ confidence level.

This result can be compared to the contours obtained from previous χ^2 fits to SNO data, shown in figure 8.9. The analysis performed here has resulted in significantly narrower contours in the $\tan^2 \theta_{12}$ dimension, equating to a better constraint on the value of this parameter. This improvement results in part from the full use of the information from all four observable parameters and the lower energy threshold of $T_{eff} = 3.5\,\text{MeV}$. The contours in $\log \Delta m_{12}^2$ are noticeably wider because the measurement of the separate day and night spectra, which was incorporated into previous SNO analyses but has not been included in this work, can be used to constrain the values in this region. Other effects that were taken into account in the previous SNO analysis include uncertainties in the shape of the predicted ^{8}B spectrum and the propagation of systematic uncertainties such as energy resolution. The addition of these effects to the analysis presented here would result in more accurate constraints on the extracted values of the mixing parameters.

8.5 Summary

A combination of improved detector modelling and reduced systematic uncertainties resulting in part from the work presented in this thesis have allowed the extraction of the neutrino fluxes and spectra to a lower energy threshold of $3.5\,\text{MeV}$. This allows for a more accurate test of possible distortions in the energy spectrum due to matter effects in the Sun or due to other new physics effects such as non-standard interactions.

This chapter has presented the flux and spectrum results from the signal extraction procedure applied in a preliminary analysis to the full LETA data set. The effects of systematic uncertainties on the observables used to characterise the data and Monte Carlo distributions have been evaluated for both an energy-constrained and unconstrained fit to the data. An evaluation of the likelihood space was used to better account for the effects of the dominant systematic uncertainties. The resulting neutrino fluxes are consistent with
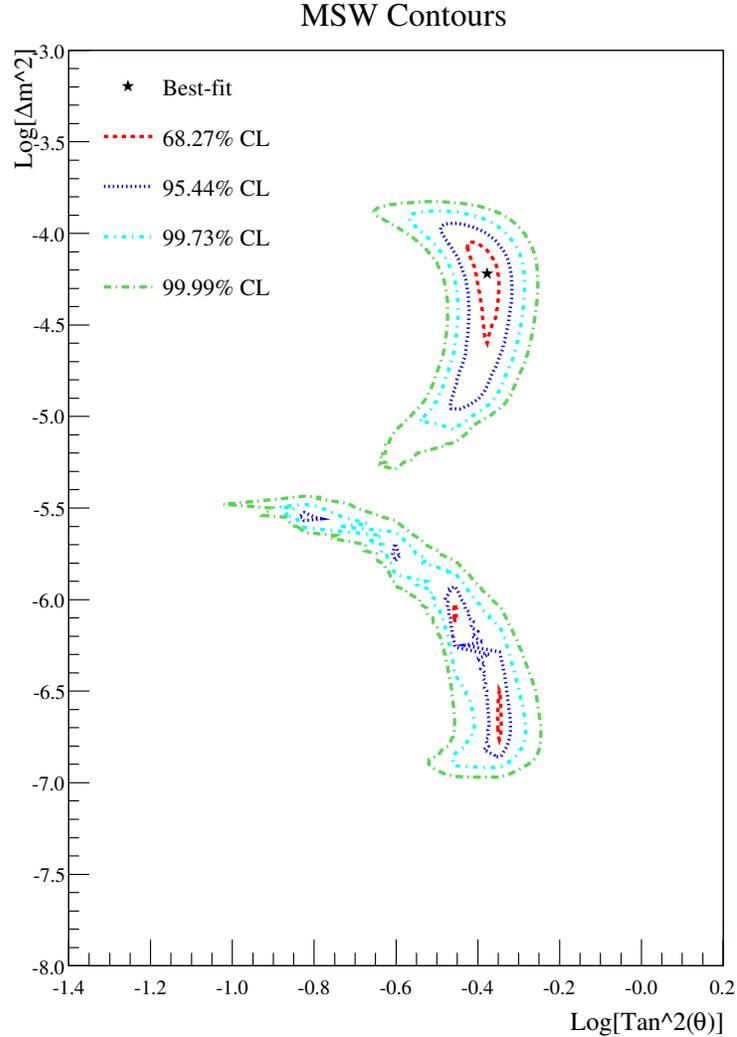
Figure 8.8: A scan of the MSW parameter space. θ and Δm^2 refer to the solar neutrino mixing parameters, θ_{12} and Δm^2_{12}. The star represents the best-fit point, defined as the point with the maximum value of the likelihood from the signal extraction. The contours show the confidence levels (CL) around this point, given by the nth sigma points ($n = 1$–4) as defined in equation 8.20.

previous results but with significantly reduced systematic uncertainties, resulting in the most accurate measurement of the solar neutrino flux ever reported. The extracted CC and ES spectra are consistent with either an undistorted spectrum or with the distortion predicted by the MSW effect.

An improved method has been investigated for utilising the full signal extraction procedure to test against MSW predictions for the neutrino fluxes and spectra. This
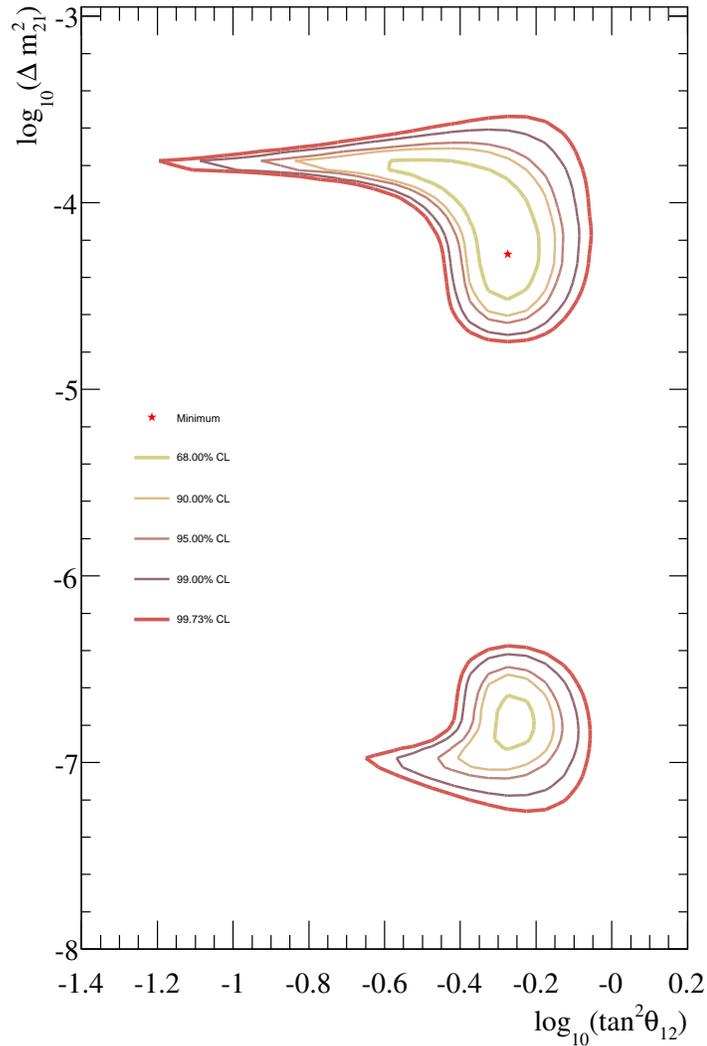
Figure 8.9: MSW contours from previous SNO analyses, using a χ^2 fit to the day and night spectra extracted from both the D_2O + salt data sets.

approach makes direct use of the likelihood calculation involved in the signal extraction, utilising the information from all the observable parameters to extract the likelihood for the predicted fluxes and spectra at each point in the MSW parameter space. Although systematic uncertainties have not yet been propagated, the resulting best-fit point is consistent with the LMA MSW prediction with reduced uncertainties in the $\tan^2 \theta_{12}$ dimension in comparison to previous SNO-only results.

Chapter 9

Conclusions

The Sudbury Neutrino Observatory has provided a clear test of the hypothesis of neutrino oscillation by observing both the total flux of ^{8}B neutrinos from the Sun and the flux of pure electron neutrinos via the NC and CC interactions, respectively. The significant deficit of ν_e observed relative to the total solar flux was conclusive proof of neutrino flavour transformation occurring en route to the Earth.

This thesis has presented an improved spectral analysis of data from the first two phases of SNO: the pure D_2O phase and the salt phase. This represents a preliminary analysis that demonstrates the techniques required and gives an indication of the expected accuracy of the final measurement, which will be performed once all remaining supporting analyses have been completed. Improved analysis techniques and detector modelling have substantially reduced the systematic uncertainties relative to previous work and allowed the extension of the analysis down to a threshold of 3.5 MeV in effective electron kinetic energy, a lower threshold than that previously reported by any water Čerenkov experiment. This new low threshold allowed a test of the predictions for possible distortions in the energy spectrum due to matter effects in the Sun.

The work described in this thesis has focused on improving the knowledge of the SNO detector in order to allow the first measurement of the solar neutrino flux and spectrum down to 3.5 MeV. A number of systematic effects were investigated in order to evaluate the associated uncertainties. In particular, any possible energy dependence was carefully studied in order to determine the impact on the final CC and ES energy spectra.

A dominant source of systematic uncertainty was the estimation of an event's energy. The algorithms used to reconstruct individual event energies rely on an accurate modelling of the optical properties of the detector, including the angular response of the light concentrators fitted to the photomultiplier tubes (PMTs). An optimisation of the modelling of this response in the simulation, presented in chapter 4, significantly improved the agreement with *in-situ* calibration data. In particular, the effects of aging were incor-

porated to account for an observed degradation in the response. As a result, a previously unexplained radial bias that had been present in the energy response of the detector was greatly reduced such that its contribution to the overall uncertainty in the energy scale was negligible.

The uncertainties associated with position reconstruction and the effect of applying a set of background rejection cuts to the data were studied in chapters 5 and 6. New analysis techniques were implemented to improve the accuracy of the measurement and to extend previous work down to the new low energy threshold. A measurement of the fiducial volume was of particular importance since it defines the number of targets available for neutrino detection and, thus, has a direct impact on the flux measurement. A bias in reconstruction causing this number to differ from reality in an energy dependent fashion would have a direct impact on the energy dependence of the flux, distorting the extracted neutrino spectrum. A mismodelling of the background rejection cuts could cause a difference between the fraction of events accepted in the data and the Monte Carlo simulation at different energies, leading to an analogous effect on the extracted neutrino spectrum. The final results significantly improved the knowledge of these effects, thus reducing the associated uncertainties.

The systematic uncertainties resulting from these and other analyses were propagated through a full signal extraction using the full LETA (Low Energy Threshold Analysis) data set, consisting of the combined D_2O and salt phase data sets. A maximum likelihood technique was used to statistically separate the individual signal types based on four observable parameters, in a three-dimensional fit for the three neutrino fluxes and 17 background event types, across the two phases of SNO data. A new technique for handling systematic uncertainties was developed to allow the data to provide additional information about the values of the uncertainties.

The total active solar neutrino flux obtained from an energy-unconstrained fit to the data was:

$$\Phi_{NC}^{uncon} = 5.191\,{}^{+0.156}_{-0.155}(\text{stat})\,{}^{+0.132}_{-0.099}(\text{syst}) \times 10^6\,\text{cm}^{-2}\,\text{s}^{-1} \qquad (9.1)$$

This is consistent with the prediction from the BS05 SSM of $5.69\pm0.85\times10^6\,\text{cm}^{-2}\,\text{s}^{-1}$ [13]. This result is consistent with previous SNO measurements but with significantly reduced uncertainties due to the improved evaluation of systematic effects, the combination of two phases of data and the lowered energy threshold. The total uncertainty on the active solar neutrino flux from this measurement is less than 4.0%, in comparison to 12.5% and 8.8% from previous SNO publications and 6.2% from a preliminary analysis of the full LETA data set down to 4 MeV. As a result, this is the most precise measurement of the

[8]B solar neutrino flux ever reported. This represents a preliminary result, with scope for even further improvement once the final analysis of the full LETA data set is performed.

Incorporating the predicted shape of the CC spectrum into the fit in order to extract a value for the CC flux, the ratio of the CC to NC fluxes was found to be:

$$\frac{\Phi_{\text{CC}}^{con}}{\Phi_{\text{NC}}^{con}} = 0.312 \, _{-0.013}^{+0.013}(\text{stat}) \, _{-0.010}^{+0.011}(\text{syst}) \tag{9.2}$$

which is consistent with previous SNO results but with significantly reduced statistical and systematic uncertainties. The observed deficit of ν_e in the solar neutrino flux, indicated by the suppression of the CC flux relative to the NC flux, is a clear indication of neutrino flavour transformation. In addition, this result is evidence for matter effects in the Sun causing further suppression of the electron neutrino survival probability since vacuum oscillation alone results in a maximum suppression of a factor of two. The CC and ES spectra were also extracted down to this lowered energy threshold and were consistent with both an undistorted spectrum and with the distortion predicted by the MSW effect.

Utilising the full likelihood evaluation from the signal extraction procedure and, thus, incorporating the information from all four observable parameters, the data was tested against model predictions in a scan of the $[\theta_{12}, \Delta m_{12}^2]$ parameter space. The best-fit point was determined to be:

$$\Delta m_{12}^2 = 6.03 \times 10^{-5} \, \text{eV}^2 \qquad \tan^2 \theta_{12} = 0.42$$

with contours that are in agreement with the LMA solution. The contours extracted from this scan are comparable to previous combined analyses of the two phases of SNO data at a higher energy threshold, but with a tighter constraint on the value of $\tan^2 \theta_{12}$. Less information is provided about the value of Δm_{12}^2 since an additional constraint on this parameter comes from the measurement of the distinct day and night fluxes, which was not incorporated into this analysis.

The work presented in this thesis has already improved the accuracy of the flux measurement over that of any previous analysis of SNO data. When the final analysis is performed, the uncertainties should be reduced even further. In particular, further scope for improvement is possible in the evaluation of the best-fit value of the energy scale and in the determination of the shape of the PMT PDF. One of the dominant limitations in this analysis was the Monte Carlo statistics available with which to build the PDFs, in particular for events originating in the AV and the PMTs. Increased statistics in these PDFs could increase the accuracy of the measurement. The resulting reduced uncertainties would allow a more stringent test of the neutrino oscillation theories, including a more sensitive search for distortions in the CC and ES spectra.

Possible extensions to this analysis include a full propagation of systematic uncertainties in the MSW model testing and incorporating a measurement of the day and night fluxes, which would place a strong constraint on the allowed value of Δm^2_{12}. The addition of information from other experiments would constrain the parameter values still further. The KamLAND reactor experiment is particularly complimentary to the SNO results, since the short baseline makes KamLAND more sensitive to the value of Δm^2_{12}, while SNO results tend to constrain the value of $\tan^2 \theta_{12}$. Therefore, combining the KamLAND measurements with the final results from the LETA data set should yield the best possible measurement of both the oscillation parameters.

Tightening the constraints on these fundamental parameters has knock-on effects on several newly developing theories, including strong implications for global θ_{13} fits. This angle has proved the most elusive due to its small value, with only an upper bound currently in place from the CHOOZ reactor experiment. An exact measurement of the value has serious implications, both for particle physics and for cosmology. In sharp contrast to mixing in the quark sector, large values have now been measured for two of the three neutrino mixing angles, θ_{12} and θ_{23}. An accurate determination of the value of θ_{13} would allow a full comparison of these two sectors of particle physics, opening up the possibility of symmetries between leptons and quarks. A measurement of θ_{13} is also the gateway to observing CP violation in the lepton sector, since a non-zero value is required for CP violation to occur. This could, in turn, answer one of the most fundamental questions of our time - the source of the matter-antimatter asymmetry in the universe. Although the value of θ_{13} has a small effect on solar neutrino oscillations, a full three neutrino analysis of the LETA data set could tighten the limit on this fundamental parameter.

Despite numerous advances in neutrino physics, the field remains full of unanswered questions. The neutrino mass hierarchy is undetermined and the absolute mass scale has yet to be measured. In addition, the very nature of the neutrino is still in question. The possibility remains open for neutrinos to be Majorana particles, which means that they are their own antiparticles. This scenario includes a massive right-handed neutrino state and implies that the origin of neutrino mass is different from that of other particles, namely, that neutrinos do not obtain mass through interactions with the Higgs particle. In addition, the interactions of Majorana neutrinos would violate lepton number conservation. Many new experiments aim to resolve this open question by searching for neutrinoless double beta decay processes, which can only occur if neutrinos are Majorana particles. In addition this would give some handle on the actual values of the neutrino masses. With such questions remaining and new ones waiting to be uncovered, the field of neutrino physics offers the possibility for many more exciting discoveries over the coming years.

Appendix A

Radioactive Decay Chains of ^{232}Th and ^{238}U

Figure A.1: The decay chain of ^{232}Th. All half-lives are shown, with the Q-values of beta and alpha decays in MeV, and gamma rays in keV. Figure courtesy of H.M. O'Keeffe [46].

^{238}U

$4.47 \times 10^9 y$
$\alpha\ 4.27$

^{234}Th

$24.10d$
$\beta^-\ 0.26\ \ \gamma 63, 92$

$m234$Pa

$1.18m$
$\beta^-\ 2.21$

^{234}U

$2.45 \times 10^5 y$
$\alpha\ 4.86$

^{230}Th

$8.0 \times 10^4 y$
$\alpha\ 4.77$

^{226}Ra

$1599y$
$\alpha\ 4.87$

^{222}Rn

$3.82d$
$\alpha\ 5.59$

^{218}Po

$3.05m$
$\alpha\ 6.11$

^{214}Pb

$26.8m$
$\beta^-\ 1.02\ \ \gamma 352$

^{214}Bi

$19.7m$
$0.021\%\ \ \alpha\ 5.62$

$> 99\%\ \ \beta^-\ 3.27$
$\gamma 609, \sim 1\%$ BR to $\gamma > 2200$

^{210}Tl

^{214}Po

$1.30m\ \ \beta^-\ 5.49$
$\gamma 800, \sim 1\%$ BR to $\gamma > 2200$

$164\mu s$
$\alpha\ 7.83$

^{210}Pb

$22.26y$
$\beta^-\ 0.06$

^{210}Bi

$5.01d$
$\beta^-\ 1.16$

^{210}Po

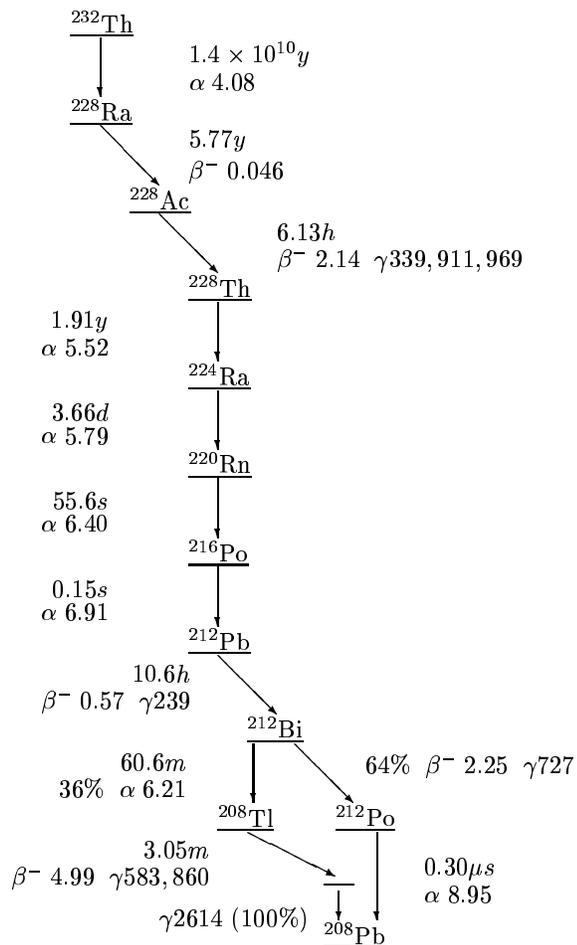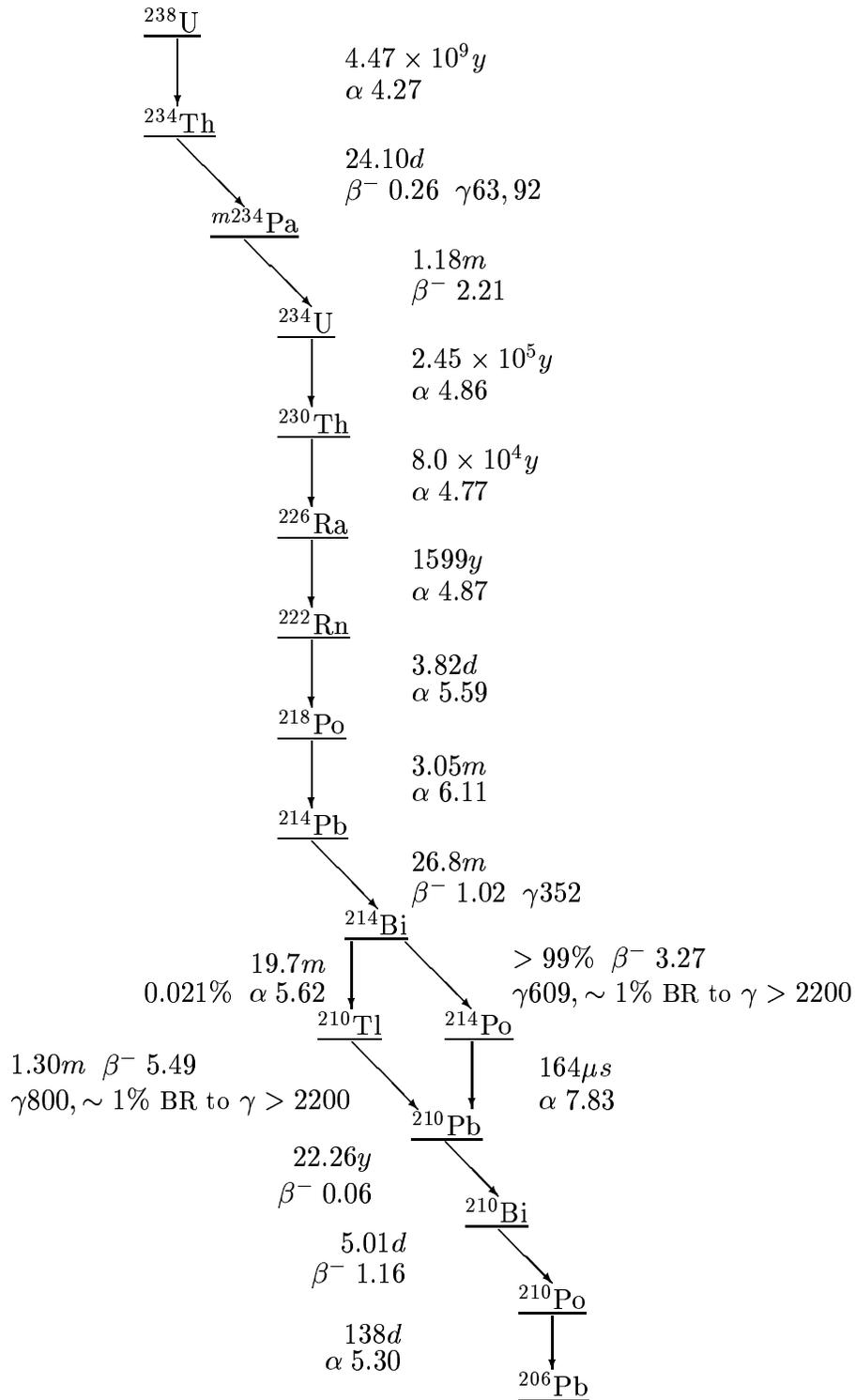$138d$
$\alpha\ 5.30$

^{206}Pb

Figure A.2: The decay chain of ^{238}U. All half-lives are shown, with the Q-values of beta and alpha decays in MeV, and gamma rays in keV. Figure courtesy of H.M. O'Keeffe [46].

Appendix B

Salt Phase Angular Response Results

The optimisation of the modelling of the PMT and light concentrator angular response in SNOMAN was discussed in chapter 4, where results were presented for the D_2O phase. The salt phase results are presented here.

The angular response in the salt phase was dealt with in three discrete time periods, defined in table 4.3. The six-wavelength fit was performed to calibration data for each of the three selected scans. Table B.1 shows the best fit values for the modified collection parameters extracted from the six wavelength fit for each scan. The change in the MCE function itself is shown in figure B.1.

Code version	a_1	a_2	a_3	b_1	b_2
Old	0.48	0.43	0.95	-0.0175	-0.007
Sep 01	5.6310	5.8012	1.0199	-0.0004	-0.0138
May 02	5.9348	6.0968	0.9978	-0.0043	-0.0553
Apr 03	5.7000	5.9012	1.0353	0.0000	-0.0353

Table B.1: The values of the parameters used in the modified collection efficiency function for the salt phase angular response. 'Old' refers to the original form of the function, before the improvements presented here.

The improvement in the fit of the Monte Carlo predicted response to the data is shown in figures B.2 - B.4 for the three optics scans. The reduced χ^2 values for the fit from 0-40° are given in table B.2, along with a comparison to the original version of the simulation. The fit is a significant improvement for all the central wavelengths (365, 386, 420, 500nm) in all three optics scans. The fit at 620nm isn't significantly different in the optimised code but, as discussed in chapter 4, the PMT sensitivity to this high wavelength is nearly zero. The fits at 337nm are all slightly worse in the new code but, again, the
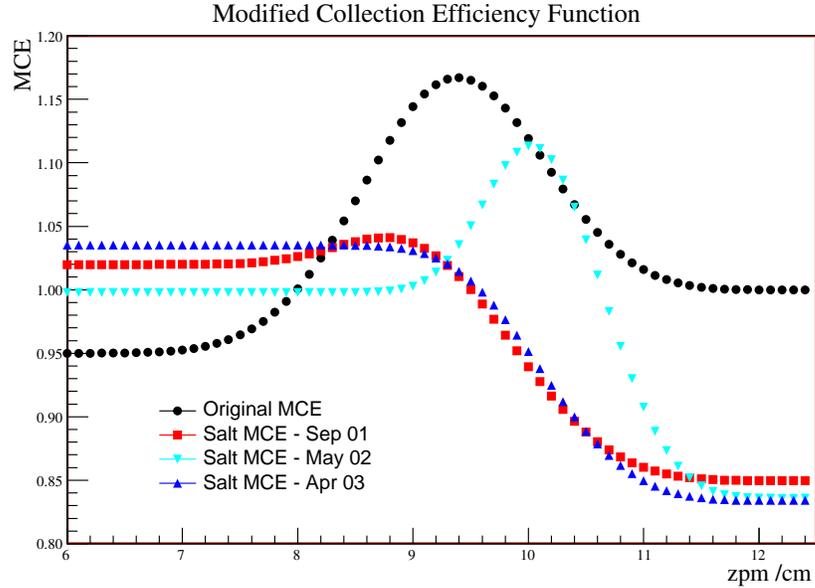
Figure B.1: The modified collection efficiency as a function of the height at which a photon strikes the PMT, for the original and three new sets of the parameters in the salt phase.

PMT sensitivity is dropping off at this value, so the slightly worsened response here is far outweighed by the improvements at other wavelengths.

Time period		Wavelength /nm					
		337	365	386	420	500	620
Sep 01	χ^2 for old MCE	18.5	147	73.8	96.1	159	715
	χ^2 for new MCE	93.9	2.70	1.42	4.30	42.4	682
May 02	χ^2 for old MCE	12.0	40.2	21.1	21.6	27.1	231
	χ^2 for new MCE	24.7	0.31	1.02	0.55	12.0	236
Apr 03	χ^2 for old MCE	13.0	92.6	36.6	54.7	94.4	535
	χ^2 for new MCE	61.0	2.24	0.57	2.64	48.7	557

Table B.2: The reduced χ^2 values for the fit of Monte Carlo prediction to calibration data for each wavelength.

A side-effect of the discrete time binning for the salt phase response was that it highlighted a heretofore unnoticed discrepancy in the normalisations between individual optics fits [84]. This could now be accounted for in the global collection efficiency, which was adjusted individually for each time bin by M. Dunford [84].

Figure B.2: A comparison of Monte Carlo prediction to calibration data for the original and the optimised versions of the simulation code for the September 2001 salt phase scan.

Figure B.3: A comparison of Monte Carlo prediction to calibration data for the original and the optimised versions of the simulation code for the May 2002 salt phase scan.
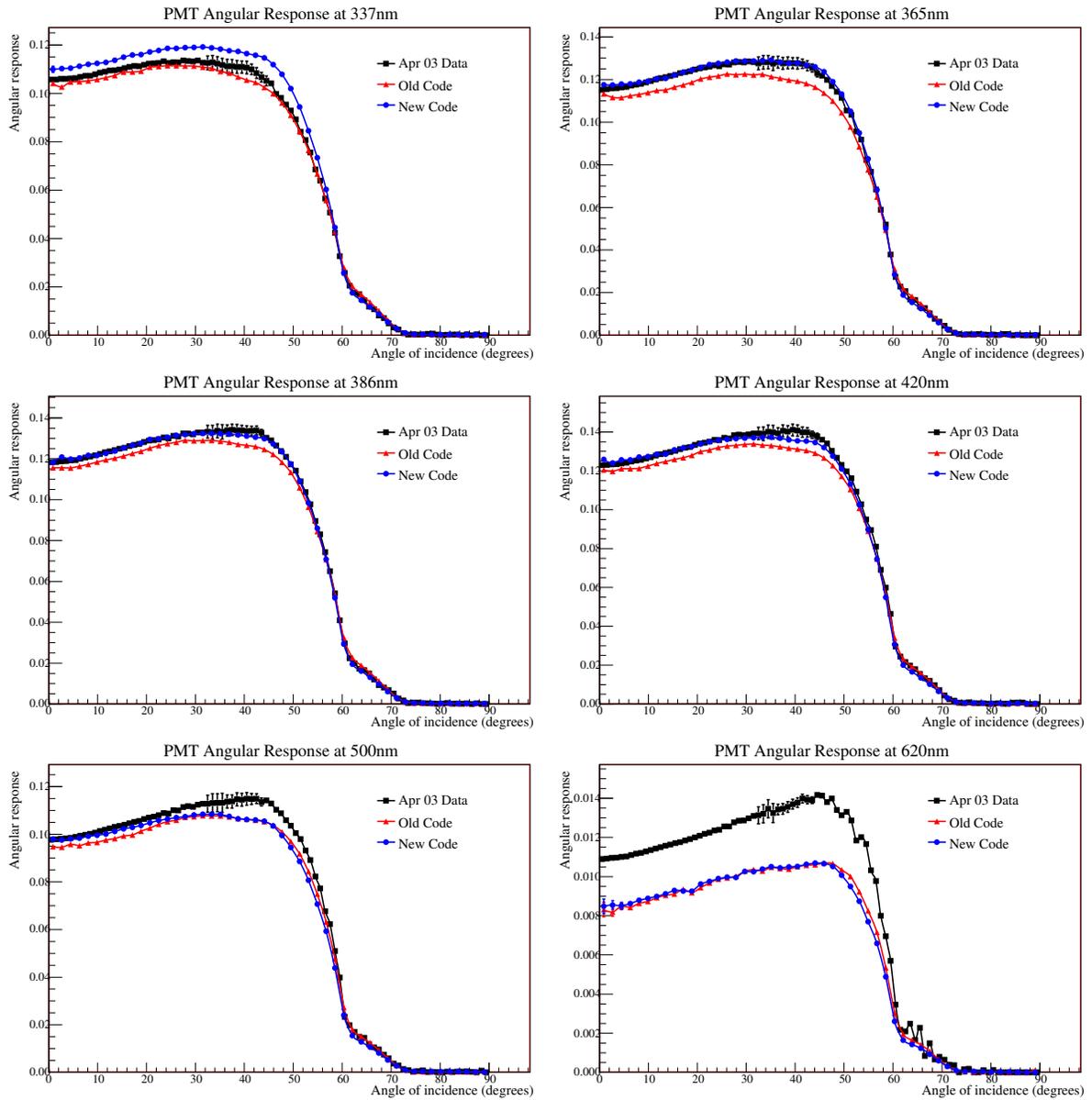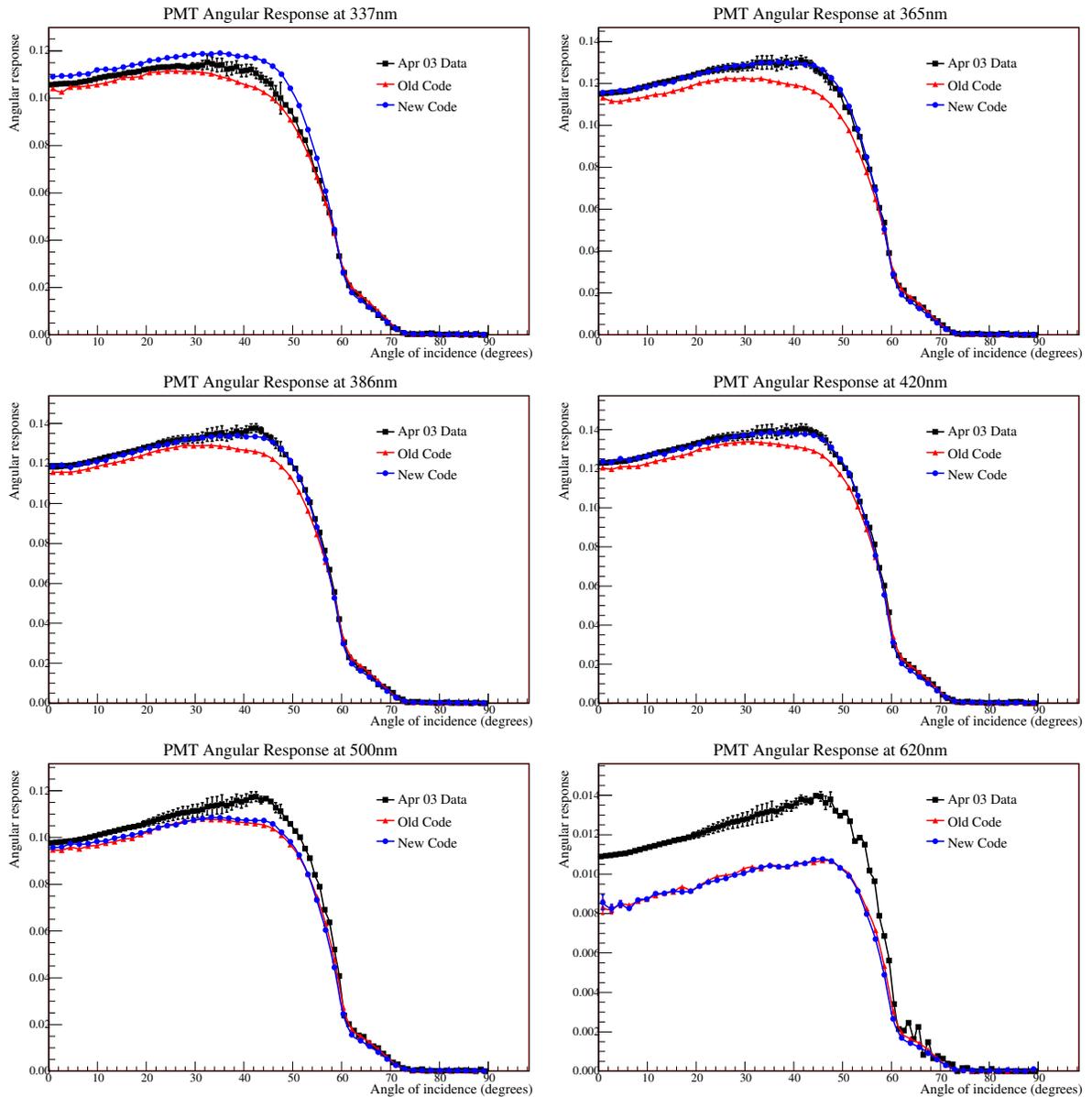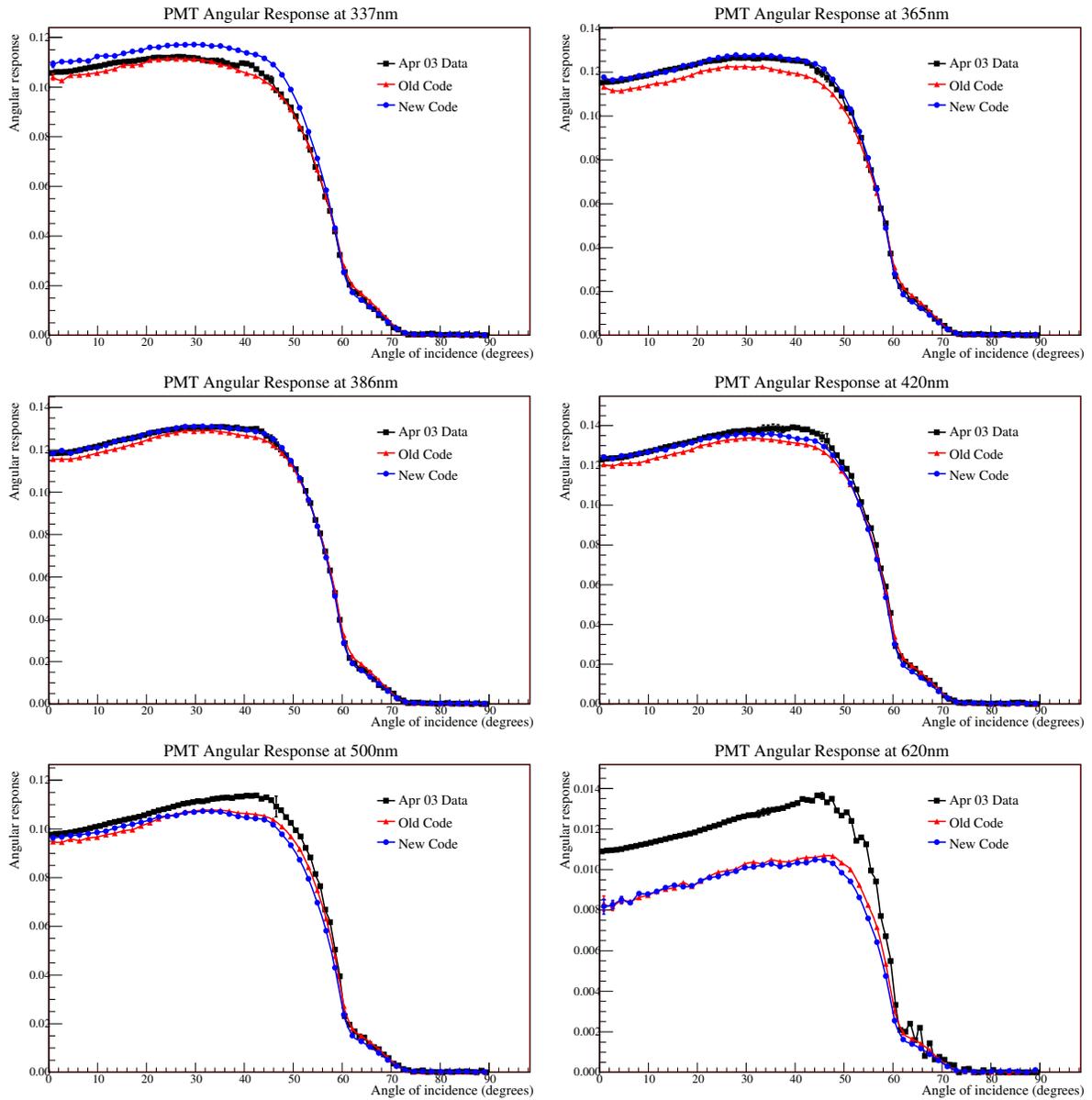
Figure B.4: A comparison of Monte Carlo prediction to calibration data for the original and the optimised versions of the simulation code for the April 2003 salt phase scan.

Appendix C

Source Geometry in the Monte Carlo Code

Both the uranium and thorium sources were deployed in the detector as encapsulated acrylic sources. In addition, a Delrin can was sometimes added around the source. The former runs are referred to as 'uncanned' and the latter as 'canned'. The dominant effect of the can was to absorb the Čerenkov light generated in the acrylic by βs emitted during the decay process.

When calibration runs were processed, the Monte Carlo simulation of all the source runs in the D_2O volume were run without modelling this can. In the H_2O region, no source geometry was simulated at all. For uncanned calibration runs this is a negligible effect and the data and Monte Carlo events can be reliably compared. However, for the canned runs, this makes a significant difference to the distributions and, so, modifications were required before the Monte Carlo simulations of these runs could be reliably used in comparison to the data. These modifications are detailed below and were utilised for the verification of background PDFs as described in chapter 7 and S. Seibert's work in [47].

C.1   Defining the Problem

For the purposes of this analysis, two thorium calibration runs were chosen to test the impact of the modifications. The runs chosen were runs 25773 and 25779 in the salt phase of data taking, with the source at $(-145, 0, -577)$. For these runs, the thorium source was deployed canned and, so, the original Monte Carlo distributions did not reproduce the data accurately, as illustrated in figure C.1. This shows the difference in the energy, β_{14} and radial distributions in the LETA analysis window for data and Monte Carlo events for one of these two runs.

These differences can be understood by considering the effect of the can. By absorbing βs emitted during the decay, the can reduces the number of generated particles

per event. A less isotropic hit pattern would therefore be expected, as is illustrated in figure C.1 by the higher value of the isotropy parameter, β_{14}.

The neutron events seen from thorium decays are the result of γs photodisintegrating a deuteron. This process would be relatively unaffected by the addition of the can around the source. However, by absorbing the β emitted during the decay, the can would reduce the number of Čerenkov events observed. The fraction of neutron events in the spectrum would therefore be higher for the canned source than the uncanned, as is observed in the energy spectra in figure C.1.

The difference in the proportion of neutron and Čerenkov events can also explain the difference in the radial distributions. Neutrons wander some finite distance before capturing whereas the β events from the uncanned Monte Carlo source can generate Čerenkov light directly adjacent to the source. The lower proportion of neutron events in the uncanned Monte Carlo simulation therefore results in a steeper radial distribution of events.
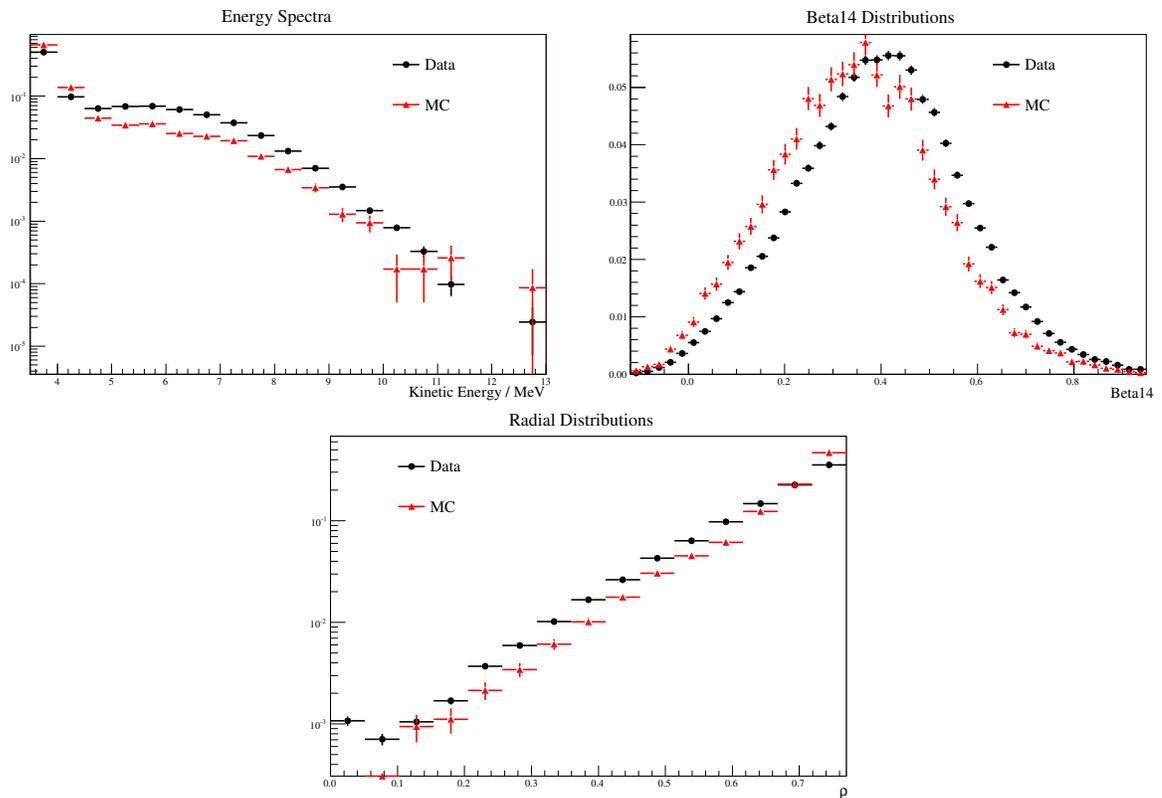

Figure C.1: Comparison of data and Monte Carlo distributions for a canned thorium run (run 25779) in the salt phase.

C.2 Modifications

For runs in the D$_2$O volume, the geometry code was modified to incorporate the effect of the source can. The simulation could then be rerun to make a fair comparison to the data. The results of this modification demonstrated that enabling the can fixed the discrepancies observed between data and Monte Carlo distributions for canned calibration runs.

For runs in the H$_2$O volume, the source geometry was not enabled due to the nature of modelling such a complicated geometry near the PSUP region. Instead, canned thorium runs were modelled by generating just the dominant γ of the decay, under the assumption that the β would be absorbed by the can. A comparison of the three versions of the geometry was performed for a thorium run in the D$_2$O volume to verify the validity of this assumption. Figure C.2 shows the comparison of the resulting Monte Carlo energy and radial distributions for each of the three versions of the geometry.

Figure C.2: Comparison of Monte Carlo distributions with different source geometries for a canned thorium run (run 25779) in the salt phase.

For illustrative purposes, the energy spectra have been fit with an exponential plus a Gaussian to represent the Čerenkov tail with the neutron peak, and the radial distributions have been fit with an exponential. For the energy and radial distributions, modelling the canned run with just the γ decay is a very good approximation. In particular, it can be seen from the fit results in table C.1 that the proportion of neutron and Čerenkov events in the spectrum is in very good agreement for the canned source and the γ approximation.

Although the β_{14} distribution for this run was well modelled by the γ approximation, when a second run was checked to confirm these results a slight offset was observed, as shown in figure C.3. The β_{14} distributions have been fit with a Gaussian. The γ simulation results in a slightly more isotropic distribution of events than the canned source

geometry, however, the approximation is still an improvement over the uncanned source. The difference may be due to source shadowing effects, which would block some of the backward going light and, hence, make the distribution for the canned source geometry less isotropic than the γ approximation. Some attenuation of the γ by the can could also result in a slightly lower energy and hence a smaller chance of multiple scattering. This would again cause the canned source to have a slightly less isotropic distribution, as observed. However, the statistics of the comparison are not high enough to draw any concrete conclusions.

The fit results for each observable for the simulation run with each of the three versions of the source geometry are given in tables C.1 - C.3. In addition, the χ^2 values for the fit of each version of the simulation to the data are given in table C.4. It can be seen that, while the uncanned Monte Carlo distributions are a very poor fit, both the canned Monte Carlo and the γ approximation reproduce the data very closely.
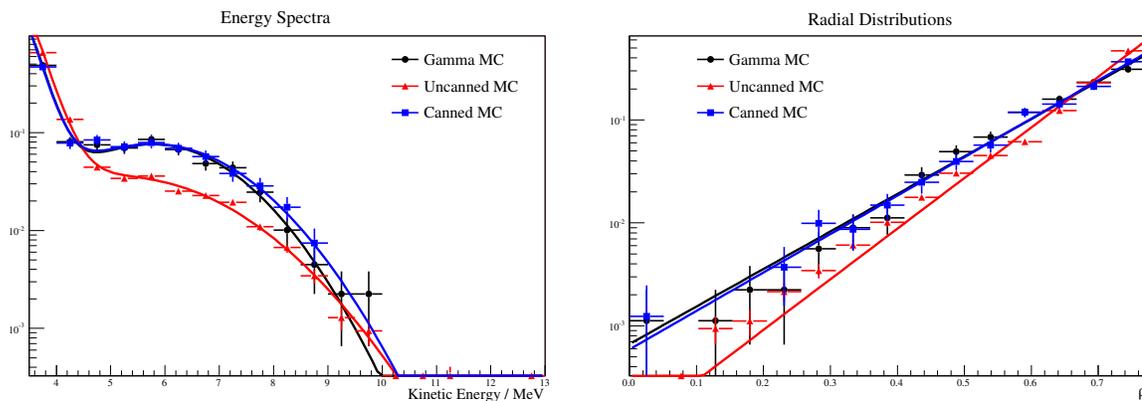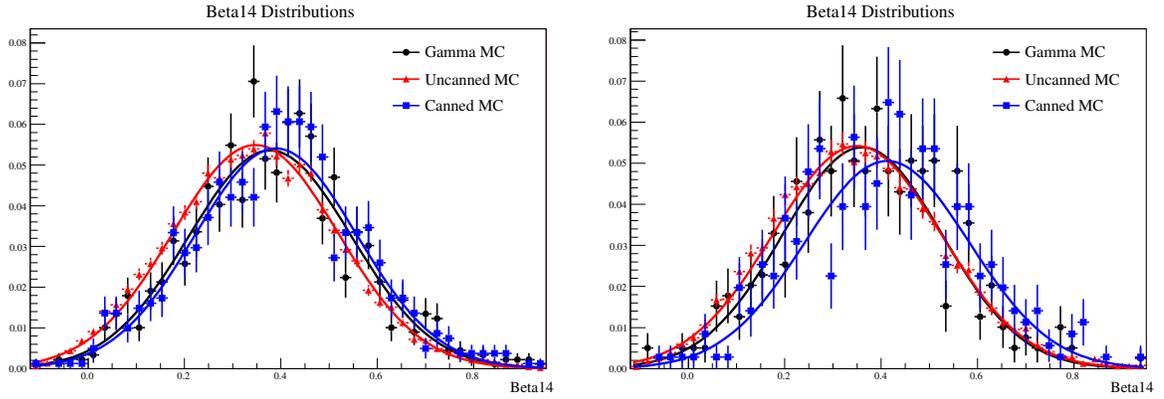

Figure C.3: Comparison of Monte Carlo β_{14} distributions with different source geometries for canned thorium runs in the salt phase. The runs used were 25779 (left) and 25773 (right).

	χ^2/dof	GNorm	Mean	Width	ENorm	Exponent
Uncanned MC	1.85	0.03±0.00	5.63±0.11	1.48±0.05	(2.7±0.8)e6	-3.5±0.1
Canned MC	1.55	0.08±0.01	5.78±0.13	1.37±0.11	(1.4±0.0)e6	-4.0±0.0
γ MC	1.25	0.08±0.01	5.78±0.11	1.26±0.08	(1.4±0.0)e6	-4.0±0.0

Table C.1: Results of fits to energy spectra for a canned thorium run (run 25779) in the salt phase. GNorm and ENorm are the normalisations of the Gaussian and exponential parts of the fit function.

	χ^2/dof	α_0	α_1
Uncanned MC	12.8	11.33±0.13	-9.26±0.09
Canned MC	0.72	8.60 ±0.34	-7.43±0.23
γ MC	1.51	8.39 ±0.26	-7.31±0.17

Table C.2: Results of exponential fits to the radial distributions for a canned thorium run (run 25779) in the salt phase. The function used was of the form $\exp(\alpha_0{}^*\rho + \alpha_1)$.

	Run 25773			Run 25779		
	χ^2/dof	Mean	Width	χ^2/dof	Mean	Width
Uncanned MC	1.67	0.35±0.00	0.17±0.00	2.96	0.35±0.00	0.17±0.00
Canned MC	0.92	0.41±0.01	0.17±0.01	0.92	0.39±0.01	0.17±0.01
γ MC	0.91	0.36±0.01	0.17±0.01	1.17	0.38±0.01	0.17±0.01

Table C.3: Results of Gaussian fits to the β_{14} distributions for canned thorium runs in the salt phase.

Observable	Monte Carlo	χ^2/dof	
	source geometry	Run 25773	Run 25779
Energy	Uncanned	159 (7)	218 (10)
	Canned	0.80 (6)	2.65 (7)
	Single γ	0.45 (7)	1.09 (7)
β_{14}	Uncanned	18.2 (24)	36.5 (24)
	Canned	1.12 (23)	1.58 (24)
	Single γ	1.21 (21)	1.69 (24)
ρ	Uncanned	37.7 (12)	62.7 (13)
	Canned	1.39 (10)	0.61 (11)
	Single γ	0.65 (13)	1.12 (13)

Table C.4: χ^2 values for the fit of the Monte Carlo distribution, run with each version of the source geometry, to the data, for canned thorium runs. The χ^2 values are quoted per degree of freedom, with the number of degrees of freedom given afterwards in parentheses.

C.3   The Uranium Source

The comparisons presented were all performed using thorium source runs. For uranium runs in the D_2O volume the can can be enabled in the Monte Carlo simulation in the same manner and the results were shown to reproduce the data to a similar level of accuracy. The uranium source was predominantly deployed uncanned in the H_2O region and so the approximation to a canned source in this region was not required.

C.4 Conclusions

For source runs that were deployed with the Delrin can, the uncanned Monte Carlo events did not reproduce the observable distributions accurately. This was due to the can absorbing the β emitted in the decay process and hence altering the nature of the event. For runs in the D_2O volume, the simulation was re-run with the can incorporated into the simulation of the source geometry and the resulting distributions reproduced the data very closely. For runs in the H_2O volume, the source geometry was not enabled but generating just the γ branch of the decay was shown to be a good approximation to a canned source run for the thorium source. The only proviso is that a small offset has been observed in the β_{14} distribution for some runs modelled in this way but the effect of this offset is negligible.

Appendix D

Using QPDT to Select PMT Events

The Monte Carlo simulation was used to generate the PDF for PMT β-γ events that was used in the signal extraction. No suitable calibration source existed with which to model these events and, so, a different method was required to verify the simulation. One of the high level cuts used to reject background events from the data set, the QPDT cut, was designed specifically to target PMT events and so could be used to provide a sample of such events against which the Monte Carlo PDF could be verified. The comparison of these 'QPDT-selected' data events to the Monte Carlo simulation is presented in chapter 7. This appendix describes the checks performed to confirm the validity of this method. In all the tests described below, every data cleaning cut utilised in the LETA analysis was applied to both data and Monte Carlo events, except the QPDT cut, which was only used to select the PMT-like events.

D.1 Selecting a Pure Sample

In order for the QPDT cut to be utilised in the verification of the PMT PDF, it was first necessary to ascertain that it provided a pure sample of PMT events. This was verified by calculating the fraction of events that were 'selected' by QPDT (failed the cut) for a number of different simulated event types. These fractions were calculated for events reconstructing inside the LETA analysis window (above $3.5\,\mathrm{MeV}$ and inside $550\,\mathrm{cm}$) and also for the lower energy thresholds of $3.25\,\mathrm{MeV}$ and $3\,\mathrm{MeV}$ to increase statistics, since the acceptance of external background events is low. These fractions are shown in table D.1.

It can clearly be seen that the QPDT-selected events inside the fiducial volume are strongly dominated by PMT events, with very low levels of contamination from radioactive decay events in either the H_2O or D_2O regions. This can be understood by the formulation of the QPDT cut: it was designed to reject events with a cluster of hits in

Event Type	Energy threshold /MeV	Fraction failing QPDT /%
PMT β-γ	3.0	13.6 ± 0.2
	3.25	14.3 ± 0.6
	3.5	14.0 ± 1.3
H_2O ^{208}Tl	3.0	0.44 ± 0.03
	3.25	0.41 ± 0.05
	3.5	0.40 ± 0.08
D_2O ^{208}Tl	3.0	0.34 ± 0.01
	3.25	0.34 ± 0.01
	3.5	0.32 ± 0.02

Table D.1: Fraction of Monte Carlo events failing QPDT for different event types.

an early time window, which is likely to identify events occurring either very close to or inside the PMTs. Although the comparison was done in terms of the fraction of events selected by QPDT, predictions of the level of each background in the data set from *ex-situ* measurements and previous analyses demonstrate that the PMT events are, in fact, the dominant background. The low fraction of other event types selected by QPDT therefore results in an even lower percentage contamination of the QPDT sample.

D.2 QPDT-Selected Monte Carlo Events

Having determined that the QPDT cut produces a relatively pure sample of PMT events, the next stage was to check the effect of the QPDT cut on the event distributions. The QPDT cut was applied in the LETA analysis and so the PMT events that were fit for during the signal extraction were those that passed this cut. Any differences in the event distributions of events passing and failing the cut must therefore be well understood in order to use the QPDT-selected events for the PDF verification.

Monte Carlo PMT events were used to make a comparison between the QPDT-selected events and those passing the cut (referred to hereafter as the 'full MC'). Lowering the energy threshold to 3.25 MeV was shown to have no noticeable impact on the shapes of the radial or β_{14} distributions for either the QPDT-selected or the full MC events and, therefore, this threshold was used for the comparison in order to improve the statistics available. Figure D.1 shows the comparison of the QPDT-selected events to the full MC for events inside the fiducial volume.

The radial and energy distributions are well reproduced, but it can be seen that the QPDT cut has a tendency to select events with higher values of β_{14}. This can again be understood by the formulation of the cut itself: events that fail QPDT are those with a
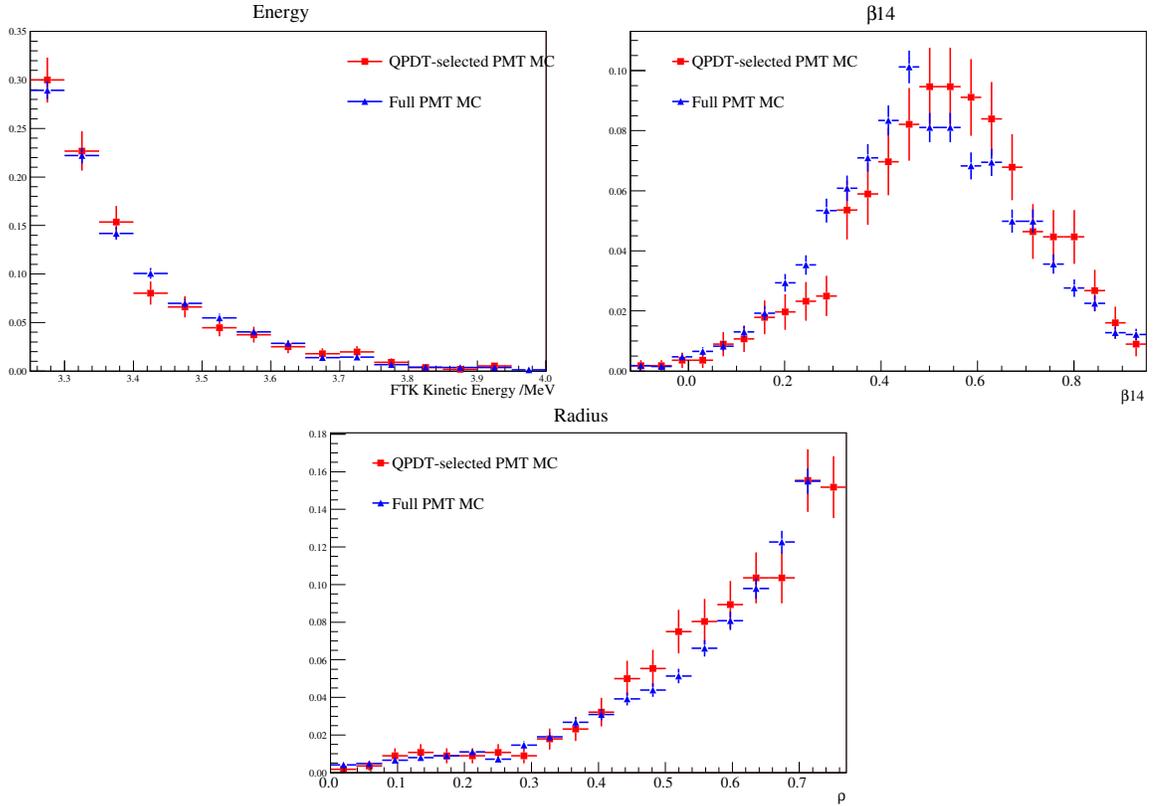
Figure D.1: Comparison of energy, β_{14} and radial distributions for QPDT-selected events and full Monte Carlo events reconstructing inside $550\,\mathrm{cm}$ and above $3.25\,\mathrm{MeV}$.

local cluster of hits far away from the reconstructed event position and, therefore, they will appear to be less isotropic than events lacking that localised hit cluster. This effect does not impact on the validity of the comparison since it would be observed for both data and Monte Carlo events.

Due to the high rate of radioactive decay in the PMTs, the data in the H_2O region is dominated by PMT events and, therefore, the data itself can be used to investigate the behaviour of such events. Since this region is outside the LETA analysis window, data events passing as well as failing QPDT can be used to confirm the effect of the cut on the observable distributions. Figure D.2 shows the β_{14} distributions for events selected by and passing QPDT in the data and the Monte Carlo simulation.

There is clear agreement between the data and Monte Carlo simulation for QPDT-selected events. The tendency of the QPDT cut to select events with higher values of β_{14} is less clear in this region of the detector, which is to be expected since the effect of the early cluster of tubes on the isotropy of an event would be less significant for an event reconstructing in the H_2O and, hence, closer to the PMTs anyway. However, the level of
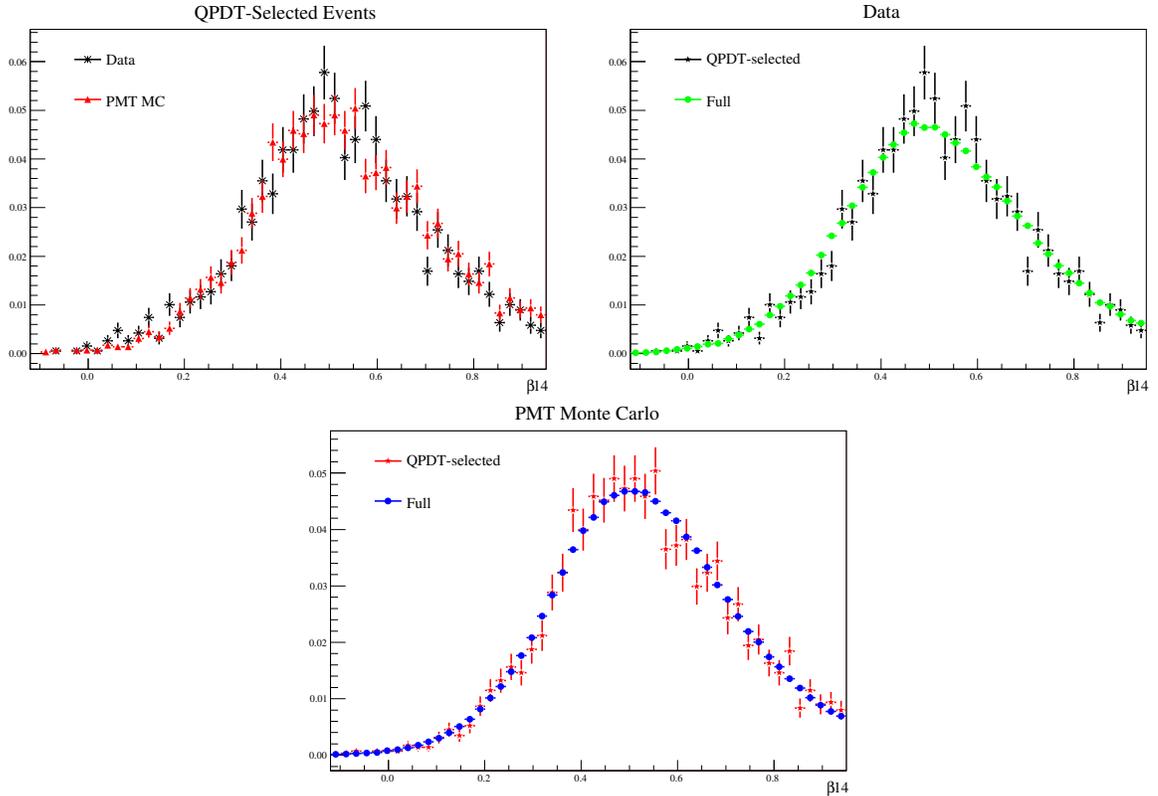
Figure D.2: β_{14} distributions for events reconstructing in the light water region. Shown are comparisons of QPDT-selected data and Monte Carlo, QPDT-selected data events to full data and QPDT-selected Monte Carlo events to full MC.

the effect is comparable for data and Monte Carlo events.

Figure D.3 shows the fraction of events selected by QPDT in the data and in the Monte Carlo PMT simulation as a function of radial position in the H$_2$O region. These fractions can be compared under the assumption that the data in the H$_2$O region is dominated by PMT events. The simulation reproduces the fraction of events selected by QPDT very accurately in the majority of this region. Near the AV ($\rho = 1.0$) a discrepancy is observed due to the breakdown of the assumption of an uncontaminated sample of PMT events in the data. The steep exponential shape of the radial distribution of PMT events means that the number of these events falls very quickly with radius and, as a result, the data near the AV contains a higher proportion of other event types. An analysis of the levels of different background types present in the H$_2$O region found that non-PMT events constitute approximately one third of the data near the AV [47]. These events would not be cut by QPDT, as demonstrated in section D.1. Therefore, the fraction of data events selected by QPDT underestimates the true fraction of PMT events selected

by this cut. Allowing for this one third contamination brings the data and Monte Carlo fractions back into very close agreement even at the AV.
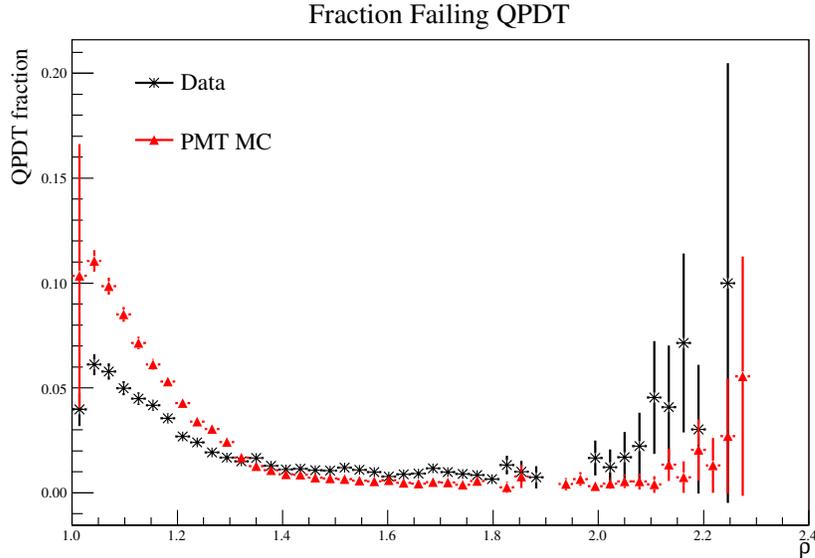

Figure D.3: The fraction of data and Monte Carlo PMT events selected by the QPDT cut in the H$_2$O region, as a function of radial position. The AV is at $\rho = 1.0$.

D.3 Summary

The QPDT cut has been shown to select a relatively pure sample of PMT events inside the fiducial volume, with very low levels of contamination from other event types. A comparison of QPDT-selected data events to similarly selected PMT Monte Carlo events in the H$_2$O region showed very good agreement. In addition, the Monte Carlo simulation predicts the fraction of PMT events removed by this cut very accurately. QPDT-selected events in the D$_2$O region have been shown to have on average higher values of β_{14} than events passing the cut, but the consistency of the behaviour of the cut on data and Monte Carlo events in the H$_2$O region is taken as evidence that this is reproducing an effect in the data. In all comparisons performed, the Monte Carlo simulation has reproduced the data very closely. Therefore, a comparison of QPDT-selected data events in the D$_2$O region to similarly selected PMT Monte Carlo events is thought to be a reliable verification of the PMT PDF.

Appendix E

Correcting for Aborted Simulation Events

The SNOMAN simulation code was written to model the SNO experiment to a high level of accuracy, including the full details of the geometry of each component of the detector. The code is therefore incredibly complex and, although it has been carefully written and thoroughly tested, rare errors have not been entirely eliminated. Occasionally, the code fails in the tracking of a photon through one of the more complicated detector regions, for example the PSUP. These 'geometry errors' cause SNOMAN to abort a small proportion of events, which artificially reduces the number of events in the Monte Carlo data set. Since this data set is used as a measure of the number of neutrino events expected inside the analysis window, a correction for these errors must be formulated. The number of photons generated in an event is directly proportional to the event's energy, so the number of such errors might be expected to exhibit an energy dependence, which must be taken into account when calculating the correction.

E.1 Method

To calculate this correction factor, Monte Carlo simulations of monoenergetic electrons were run across the energy range used for the LETA analysis in the same 0.5 MeV steps as were used for the neutrino flux analysis described in chapter 8. Samples of 500,000 events were generated at each energy and the number of aborted events in each sample was counted in order to calculate the correction for each set of events. The run conditions for each simulation, including the number and position of working PMTs, were set to those of run 20684, which was the first neutrino run in the salt phase data set. Although it would have been more accurate to calculate this correction individually for each run and weight the results over the entire livetime, this is a second order effect to a small correction and would have a negligible impact on the final result.

The method described above calculates the correction factor for Monte Carlo events in terms of the generated event energy. When applying this correction, it was applied in terms of reconstructed energy rather than generated energy, so a check was performed to verify that this approximation was valid. Running the full energy reconstruction algorithm during event generation slows the process down significantly, so 50,000 events were generated at each energy for this verification. The correction was recalculated for this reduced data set both in terms of generated kinetic energy and in terms of the average reconstructed kinetic energy for the events at each energy interval. Extra smearing of the latter result would occur due to the finite energy resolution associated with reconstruction, but this is once more a second order effect on a small correction. The results calculated in this way were used purely for the verification and not for the final correction since the statistics were more limited and the correction therefore less accurate.

E.2 Results

To correct the Monte Carlo data set to account for the aborted simulation events, each Monte Carlo event was weighted by a factor greater than or equal to 1.0. This weighting, w, was evaluated using the fraction of successfully detected events, as follows:

$$
\begin{aligned}
N_d &= N_T - N_f \\
N_T &= w \times N_d \\
w &= \frac{N_T}{N_T - N_f} = \frac{N_T}{N_d}
\end{aligned}
\tag{E.1}
$$

where N_T is the number of 'true' events in the simulation, N_f is the number that were aborted due to geometry errors, N_d is the number that were detected in the Monte Carlo data set and w is the resultant weighting factor to correct for the aborted events.

Figure E.1 shows the fraction of events detected as a function of the generated electron kinetic energy. A first order polynomial was fit to parameterise the energy dependence, resulting in the following form for the fraction of successfully detected Monte Carlo events:

$$
MC_{frac}(T_{eff}) = 1.00 - 0.0006238 \times T_{eff}
\tag{E.2}
$$

As described above, each Monte Carlo event needs to be weighted by the reciprocal of the fraction of observed events, so the weighting for each event is given by

$$
w(T_{eff}) = \frac{1}{MC_{frac}(T_{eff})}
\tag{E.3}
$$

Figure E.2 shows the same fraction calculated for the reduced data set in terms of both the generated electron kinetic energy and the average reconstructed kinetic energy. The results in the two cases agree very closely within the statistical uncertainties on the fit parameters. This supports the approximation inherent in weighting events according to their reconstructed energy using the correction that was calculated in terms of generated energy, as given in equations E.2–E.3.
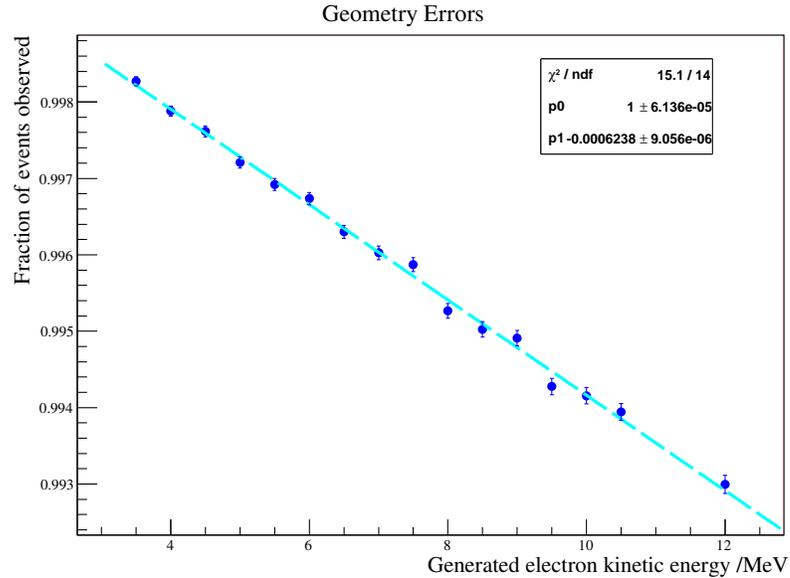

Figure E.1: Fraction of successfully detected events, accounting for aborted simulation events, in terms of the generated electron kinetic energy.
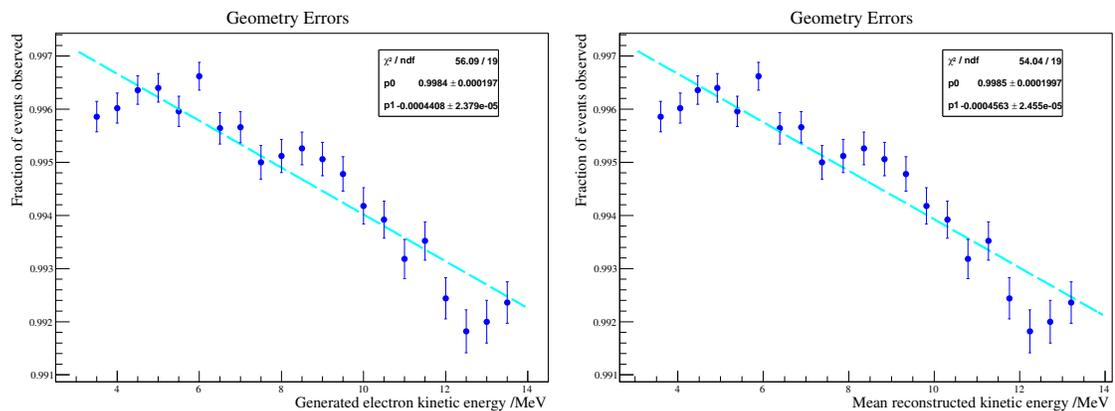

Figure E.2: Fraction of successfully detected events in terms of the generated electron kinetic energy (left) and the average reconstructed kinetic energy (right).

Appendix F

Use of the Likelihood Function to Evaluate Systematic Uncertainties

The effect of systematic differences in the observable parameters between data and the Monte Carlo simulation used to generate the PDFs must be taken into account in the signal extraction. These differences can be parameterised in terms of systematic shifts or scalings to be applied to each event.

In principle, parameters with such systematic uncertainties should be allowed to vary within those uncertainties in the context of the maximum likelihood approach, in the same way as the flux parameters. In practice, this can complicate the fit procedure owing to the large number of potential variables, many of which have a negligible impact on the final results. The approach taken here was to identify those parameters that have the largest effect and to use the likelihood from the fit to determine their true values. A manual 'scan' of the likelihood space in a particular parameter was used to verify the true best-fit value and to extract accurate values for the associated uncertainty, equating to the points at which the likelihood of the fit was reduced by 0.5 from the best-fit value. The results of these scans are presented here. The scans were performed for both the energy-constrained and unconstrained fits. The results were similar in both cases, with only the final extracted values of the uncertainties varying slightly between the two cases.

External measurements of the systematic effects provide independent information about the uncertainty in each one. For each systematic parameter, the measured values of the uncertainty from external analyses were applied as a penalty to the likelihood function in the form of a Gaussian constraint term, as described in chapter 7.

F.1   Energy Scale

The likelihood space was scanned for each of the three energy scales independently. For the correlated part of the energy scale, the energies of events in both phases were scaled

together whereas, for the uncorrelated energy scale in each phase, only the energies of events in that phase were scaled. The measured values from external analyses, given in table 8.1, were used to impose a Gaussian penalty term on the likelihood function for each scan in order to appropriately constrain the values of the energy scale.

The scan in correlated energy scale for the constrained fit revealed the existence of a new minimum at a shift of +0.75% from the central fit. This is shown in the upper left plot in figure F.1, which shows the change in the negative log likelihood (NLL) from the central fit as a function of the scaling applied to event energies. The change in the likelihood is approximately -3.0 from the unshifted point, which represents a significance of greater than 2σ. Therefore, the PDFs were corrected by this amount and the scan was repeated. The result of this second scan is shown in the upper right plot in figure F.1. The shift to the new minimum was clearly successful and, so, all further likelihood scans were performed around this new minimum.

The scan for the unconstrained fit revealed the existence of two minima, of similar levels of significance (within 1–2σ of each other). The tests performed on artificial data sets (section 7.5.4) demonstrated that fluctuations in the likelihood at this level were common due to events shifting across bin boundaries. Therefore, it is not straightforward to reliably determine which of the two locations is the true minimum. For the final analysis, both minima should be explored in more detail, for example by scanning the likelihood more finely around each one, to determine which is the true minimum. However, the locations of the minima are within 2σ of each other in the measured uncertainty on the correlated energy scale and, thus, are consistent with the knowledge of the detector. In addition, the significance of the minima (represented by their change in NLL from the central fit point) are also consistent to within 2σ. The effect of the shift in the energy scale on the neutrino fluxes and spectra is small in comparison to other uncertainties and, therefore, is not significant. Since the PDFs assume an undistorted model, for the purposes of this thesis the comparison to the constrained fit was used to select the minimum at a shift of +0.75% in order to test the null hypothesis of no distortion in the energy spectrum. Therefore, the PDFs were also corrected by this amount for the unconstrained fit. The scans both before and after this correction are shown in the lower plots in figure F.1. In order to account for the uncertainty inherent in the choice of one minimum over the other, the negative uncertainty on the correlated energy scale was inflated to -1.5% to include the value of the second minimum in the one sigma bound.

The one sigma bounds for the uncertainty on the correlated energy scale were taken as the points at which the NLL increased by 0.5 from the new central fit point. In some cases, the discrete nature of the sampling of the likelihood space resulted in some discontinuities in the likelihood values near the minimum, which could artificially reduce
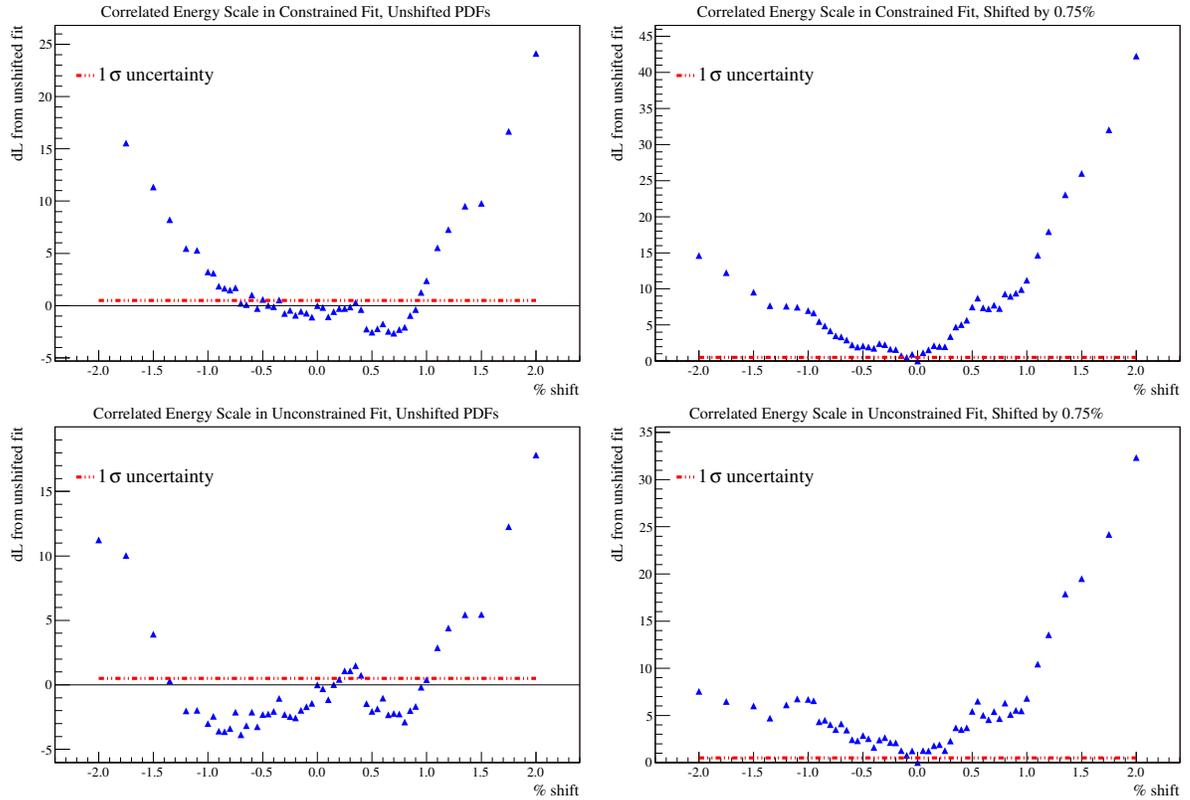
Figure F.1: Results of scanning the likelihood space in the correlated energy scale for the constrained (upper) and unconstrained (lower) fits, with no shifts (left) and a shift of $+0.75\%$ (right) applied to the PDFs. The dashed red line represents the $1\,\sigma$ bound from the central fit point, given by a change in NLL of $+0.5$.

the uncertainties extracted from the scan. In these cases, a smooth polynomial was fit in a small region around the minimum in order to conservatively determine where the NLL increased by 0.5. The resulting values are quoted in table 8.1. The final values extracted for the uncertainties differed slightly for the two fit types. This is to be expected for the energy scale in particular since the additional information from the CC and ES spectra in the constrained fit results in a tighter constraint on the value of the systematic parameter and, hence, smaller values of the uncertainty were extracted for this fit type.

The scans in the uncorrelated energy scales indicated no further new minima. Given the possible overlap between the effect of the correlated and uncorrelated energy scales, the uncorrelated scales were scanned both with and without the application of the shift discussed above. The scans without the shift applied indicated a minimum at a positive shift for the D_2O phase energy scale for both fit types. The scans with the shift applied were centred around zero, indicating that the shift applied to the correlated energy scale

also had some impact on this parameter. The one sigma bounds on the uncertainties were again taken as the points at which the NLL increased by 0.5 from the central fit, where the central fit now refers to the fit with the applied +0.75% energy scaling.

The uncertainties on the correlated and uncorrelated energy scales resulting from the likelihood scans, for both the constrained and unconstrained fits, are given in table 8.1 in chapter 8.

F.2 β_{14} Scaling and Resolution

The likelihood space was scanned for both the mean and the width of the β_{14} distribution for electron events. These uncertainties were both treated as correlated between the phases. The measured values for the uncertainties from external analyses are given in table 8.2. These values were used for the widths of the Gaussian penalty terms that were applied to the likelihood function for each scan in order to constrain the values of the parameters. Although these parameters were treated as correlated between the phases, an independent measurement was made in each phase and, therefore, two constraints were applied to the likelihood.

The scan in the β_{14} electron scale indicated a new minimum at a shift of +0.3% for both the constrained and unconstrained fits, at a change in NLL from the central fit of approximately -3.0, which represents a significance of greater than $2\,\sigma$. Therefore, the PDFs were corrected by this shift and the scan was rerun. The results of the scan both before and after the application of this shift are shown in figure F.2. These figures show the change in the NLL from the central fit as a function of the scaling applied to the β_{14} values of electron-like events.

Applying this shift to the PDFs successfully resulted in a new minimum centred on zero. Therefore, this point was used as the central fit for all further likelihood scans. The one sigma bounds for the uncertainty on the correlated β_{14} electron scale were taken as the points at which the NLL increased by 0.5 from the new central fit point. These are indicated by the dashed red line on the right hand plot in figure F.2 and the resulting values are quoted in table 8.2.

To account for possible correlations between parameters, the scans of the energy scale presented in section F.2 were rerun to determine the impact of this additional shift. No change was observed in the position of the minimum, demonstrating that the location of the minimum in one parameter was robust to changes in the other. However, the values of the bounds on the energy scale uncertainties were actually reduced by the application of the shift to the β_{14} values. Therefore, the values quoted for the uncertainties on the energy scale were those evaluated at this new minimum since this is at the point at which
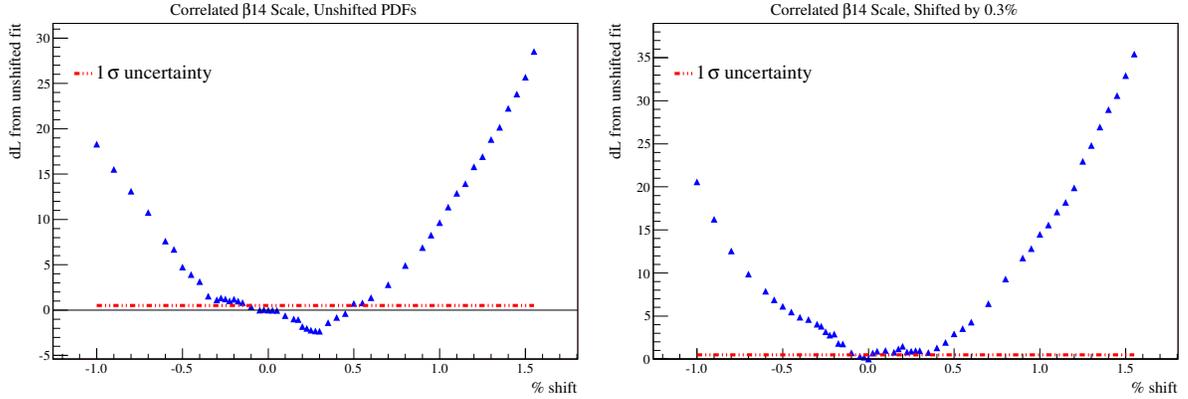
Figure F.2: Results of scanning the likelihood space in the electron β_{14} scaling systematic with no shifts (left) and a shift of $+0.3\%$ (right) applied to the PDFs. The dashed red line represents the $1\,\sigma$ bound from the central fit point, given by a change in NLL of $+0.5$. Both figures are from the scan performed using an unconstrained energy spectrum.

the final signal extraction was performed.

The scan in the electron β_{14} width was centred around zero, indicating no need for a further correction to the PDFs. Therefore, the values for the uncertainty in this parameter were taken as the point at which the NLL increased by 0.5 from the central fit, where the central fit now refers to the fit with the $+0.75\%$ energy scaling and the $+0.3\%$ electron β_{14} scaling applied. The values extracted for both the mean and width uncertainty on the electron β_{14} are given in table 8.2 in chapter 8, for both the constrained and unconstrained fits.

F.3 Energy Resolution

The likelihood space was scanned for the D_2O and salt phase resolutions independently. As before, a penalty term was applied to the likelihood according to prior knowledge of the magnitude of the systematic effect. In the case of the energy resolution, a correction had already been applied to the Monte Carlo events to account for the broader energy resolution seen in calibration data and, so, the penalty came from the uncertainty on this correction.

A new minimum was observed, in both the D_2O and salt phases, for both fit types. The significance of the minima were at a level of 2–$3\,\sigma$ in each case. The locations of the minima indicated that an additional smearing should be applied to the PDFs of $0.02\,\text{MeV}$ in the D_2O phase and $0.025\,\text{MeV}$ in the salt phase. Interestingly, when combining this smearing with the original correction described in section 8.2.1, the result is an equal

smearing of 0.17 MeV in each phase.

The same technique was applied to extract values for the uncertainties in each parameter. In some cases, the extracted uncertainties were actually larger than those measured using calibration sources. In addition, several of the resulting uncertainties were found to be quite asymmetric. However, the likelihood technique should result in the most accurate measure of the one sigma bounds on each parameter and, so, these values were propagated to the signal extraction as the uncertainties on the energy resolution. As was seen for the energy scale, the discrete nature of the sampling of the likelihood space occasionally resulted in the possibility of underestimating the values of the uncertainties. In these cases, a smooth polynomial was fit in a small region around the minimum in order to conservatively determine where the NLL increased by 0.5. The resulting bounds on the uncertainties in the energy resolution are given in table 8.1 in chapter 8. The asymmetry observed in some of the scans is illustrated in figure F.3, which shows the scan of the D_2O phase energy resolution for the unconstrained fit both before and after the application of the correction.
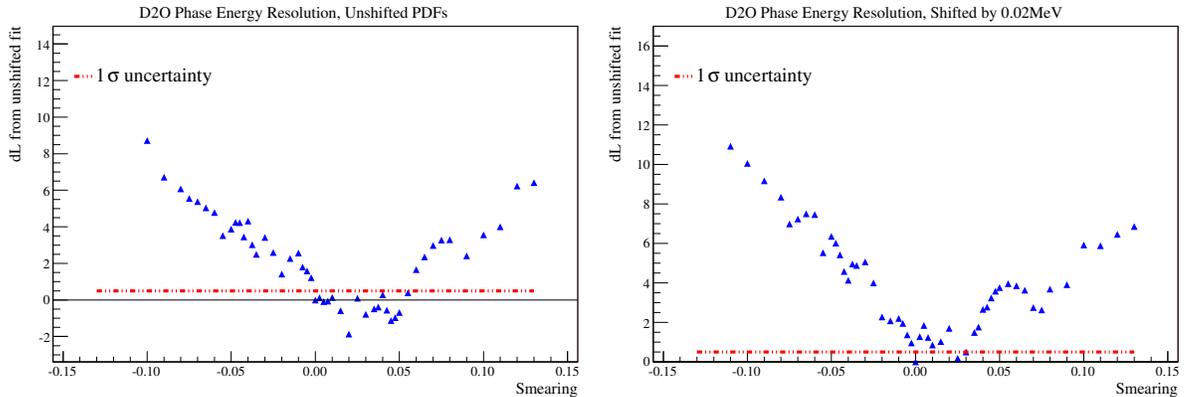

Figure F.3: Result of scanning the likelihood space in the D_2O phase energy resolution with no shifts (left) and a shift of 0.02 MeV (right) applied to the PDFs. The dashed red line represents the $1\,\sigma$ bound from the central fit point, given by a change in NLL of +0.5. Both figures are from the scan performed using an unconstrained energy spectrum.

To account for possible correlations, the scans of the energy scale and the β_{14} scale and width, presented in sections F.1 and F.2, respectively, were rerun to determine the impact of this additional smearing in the energy. No change was observed in the position of the minima found for these parameters, with each scan still resulting in a minimum centred on zero additional shift. This indicates that the effect of the energy resolution is at most weakly correlated with the other parameters and that the location of the minimum in one parameter is robust to changes in the others.

F.4   Energy-Dependent Fiducial Volume

The measurement of the energy-dependence to the fiducial volume was treated as un-correlated between the phases and the likelihood space was scanned for each phase independently for both the constrained and unconstrained fits. The measured values for the uncertainties used as the penalty terms in the likelihood functions were taken from chapter 5. In all cases, very little extra information was provided by the data. The values extracted for the uncertainties in this measurement were very close to the independent measurements used to constrain it and, so, these measurements were used as the bounds in the final signal extraction.

An example of the likelihood space is shown in figure F.4 for the salt phase, for the unconstrained fit. The bounds on the measurement extracted from this scan, given by the points at which the NLL has increased by 0.5 from the central fit, are +0.4, -0.45%. In comparison, the independently measured uncertainties on this parameter (from chapter 5) were +0.41, -0.48%.
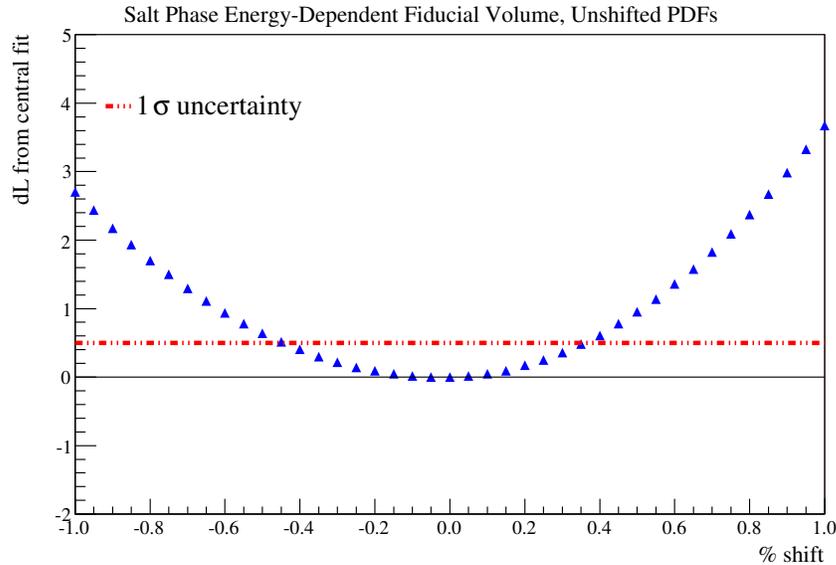

Figure F.4: Result of scanning the likelihood space for systematic uncertainties in the energy-dependent fiducial volume measurement in the salt phase, using an unconstrained energy spectrum.

Bibliography

[1] Laurie M. Brown. "The Idea of the Neutrino." *Physics Today*, **31**, 27 (1978). Includes an English translation of W. Pauli's December, 1930 letter.

[2] F. Reines & C.L. Cowan. *Nature*, **178**, 446-9 (1956).

[3] K. Hagiwara *et al. Phys. Rev. D*, **66**, 010001 (2002).

[4] V. Gribov & B. Pontecorvo. *Phys. Lett. B*, **28**, 493-6 (1969).

[5] L. Wolfenstein. *Phys. Rev. D*, **17**, 2369 (1978).

[6] S.P. Mikheyev & A.Y. Smirnov. *Sov. J. Nucl. Phys.*, **42**, 913 (1985).

[7] Super-Kamiokande Collaboration. *Phys. Lett. B*, **433**, 9-18 (1998).

[8] Super-Kamiokande Collaboration. *Phys. Rev. D*, **71**, 112005 (2005).

[9] K2K Collaboration. *Phys. Rev. Lett.*, **94**, 081802 (2005).

[10] MINOS Collaboration (2007). ArXiv:hep-ex/0708.1495v2.

[11] CHOOZ Collaboration. *Eur. Phys. J. C*, **27**, 331-374 (2003).

[12] SNO collaboration. *Phys. Rev. C*, **72**, 055502 (2005).

[13] John N. Bahcall *et al. ApJ*, **621**, L85 (2005). ArXiv:astro-ph/0412440v3.

[14] John Bahcall. "http://www.sns.ias.edu/ jnb".

[15] R. Davies Jr *et al. Phys. Rev. Lett.*, **20**, 1205-9 (1968).

[16] B.T. Cleveland *et al. ApJ*, **496**, 505 (1998).

[17] P. Anselmann *et al. Nuc. Phys. B (Proc. Suppl.)*, **31**, 117 (1993).

[18] M. Altmann *et al. Phys. Lett. B*, **490**, 16 (2000).

[19] J.N. Abdurashitov *et al. Nucl. Phys. Proc. Suppl.*, **118**, 39 (2003).

[20] S. Fukuda *et al. Phys. Rev. Lett.*, **77**, 1683-6 (1996).

[21] Super-Kamiokande Collaboration (2008). ArXiv:hep-ex/0803.4312v1.

[22] Mark Thomson. "Neutrino Oscillations." *University of Cambridge (published online)*. (2002).

[23] H.H. Chen. *Phys. Rev. Lett.*, **55**, 1534-6 (1985).

[24] SNO Collaboration. *Phys. Rev. Lett.*, **87**, 071301 (2001).

[25] SNO Collaboration. *Phys. Rev. Lett.*, **89**, 011301 (2002).

[26] KamLAND Collaboration. *Phys. Rev. Lett.*, **94**, 081801 (2005).

[27] Borexino Collaboration (2008). ArXiv:astro-ph/0708.2251v2.

[28] O.G. Miranda *et al. JHEP*, **0610**, 008 (2006). ArXiv:hep-ph/0406280v3.

[29] V. Barger *et al. Phys. Rev. Lett.*, **95**, 211802 (2005). ArXiv:hep-ph/0502196v2.

[30] The SNO Collaboration. *Nuclear Instruments & Methods A*, **449**, 172-207 (2000).

[31] J.R. Klein *et al.* "The SNO Trigger System." (1997). SNO Technical report SNO-STR-97-035, University of Pennsylvania.

[32] Stan Seibert & Christopher Tunnell. "Measuring the Trigger Efficiency in the NCD-Phase and the All-Phase Implications." (2006). Technical report, University of Texas at Austin.

[33] T.C. Anderson *et al.* (SNO). *Nucl. Instrum. Meth.*, **A501**, 399 (2003).

[34] T.C. Anderson *et al.* (SNO). *Nucl. Instrum. Meth.*, **A501**, 386 (2003).

[35] Charlotte Sims. "Background Determination in the Salt Phase of the Sudbury Neutrino Observatory Experiment." (2005). PhD thesis, University of Oxford.

[36] Jeanne Wilson. "A Measurement of the ^{8}B Solar Neutrino Energy Spectrum at the Sudbury Neutrino Observatory." (2004). PhD thesis, University of Oxford.

[37] M.R. Dragowsky *et al. Nuclear Instruments & Methods A*, **481**, 284-296 (2002).

[38] N. Tagg *et al. Nuclear Instruments & Methods A*, **489**, 92-102 (2002).

[39] Mark Kos. "Low Energy Solar Neutrino Analysis of the Salt Phase of the Sudbury Neutrino Observatory." (2007). PhD thesis, Queen's University.

[40] The SNO Collaboration. "The SNOMAN Companion." Ed. 5.03.

[41] CERN Program Library Long Write Up. "The Zebra System." (1995). Q100/Q101.

[42] CERN Program Library Long Write Up. "HEPDB - Database Management Package." (1995). Q180.

[43] W.R. Nelson *et al.* "The EGS4 Code System." (1985). Stanford Linear Accelerator Centre Report-265.

[44] Los Alamos National Laboratory. "MCNP - A Monte Carlo N-Particle Transport Code System." (1993). V4c.

[45] Jessica A. Dunmore. "The Separation of CC and NC Events in the Sudbury Neutrino Observatory." (2004). PhD thesis, University of Oxford.

[46] Helen M. O'Keeffe. "Low Energy Background in the NCD Phase of the Sudbury Neutrino Observatory." (2008). PhD thesis, University of Oxford.

[47] SNO Collaboration. "LETA Unidoc." (2008). Technical report, Sudbury Neutrino Observatory.

[48] SNO collaboration. *ApJ*, **653**, 1545 (2006).

[49] SNO collaboration. *Phys. Rev. C*, **75**, 045502 (2007).

[50] M. Dunford. "Normalizing the PMT Reflections." (2005). Technical report, University of Pennsylvania.

[51] Monica Dunford. "Measurement of the ^{8}B Solar Neutrino Energy Spectrum at the Sudbury Neutrino Observatory." (2006). PhD thesis, University of Pennsylvania.

[52] Josh Klein & Mark Neubauer. "Using Time and Angle Information in Reconstruction and the SNOMAN Path Fitter." (2001). Technical report, University of Pennsylvania.

[53] Mark Guy Boulay. "Direct Evidence for Weak Flavour Mixing with the Sudbury Neutrino Observatory." (2001). PhD thesis, Queen's University.

[54] Neil K. McCauley. "Producing a Background Free Data Set for Measurement of the Charge Current Flux and Day-Night Asymmetry at the Sudbury Neutrino Observatory." (2001). PhD thesis, University of Oxford.

[55] Aksel Hallin. "^{24}Na in the Neck, III." (2003). Technical report, Sudbury Neutrino Observatory.

[56] Martin Emilio Moorhead. "Reflectors in Cherenkov Detectors." (1992). PhD thesis, University of Oxford.

[57] Michael David Lay. "Creation and Detection of Čerenkov Light in the Sudbury Neutrino Observatory." (1994). PhD thesis, University of Oxford.

[58] G. Doucas *et al. Nuclear Instruments & Methods A*, **370**, 579-596 (1996).

[59] Matthew Jeremy Lyon. "Neutron Transport in the Sudbury Neutrino Detector." (1996). PhD thesis, University of Oxford.

[60] H.M. O'Keeffe. "The dependence on aging of the angular response of a SNO concentrator." (2003). Internal report, Oxford University.

[61] H.M. O'Keeffe. "Improved Measurements of SNO Reflector Response." (2004). Internal report, Oxford University.

[62] Aksel Hallin. "SNOMAN 3D PMT." (2003). Technical report, Sudbury Neutrino Observatory.

[63] Robert James Boardman. "The Detection of Čerenkov Radiation from Neutrino Interactions." (1992). PhD thesis, University of Oxford.

[64] S.D. Biller *et al.* "Optics 6b." (2001). Technical report, Sudbury Neutrino Observatory.

[65] O. Simard. *Private Communication.*

[66] CERN Program Library Long Writeup D506. "MINUIT: Function minimization and error analysis."

[67] M. Dunford. "Studies of Wavelength Dependent Angular Response in Salt." (2006). Technical report, University of Pennsylvania.

[68] Aksel Hallin (2007). *Private Communication.*

[69] Mark Stephen Neubauer. "Evidence for Electron Neutrino Flavour Change Through Measurement of the ^{8}B Solar Neutrino Flux at the Sudbury Neutrino Observatory." (2001). PhD thesis, University of Pennsylvania.

[70] Kathryn Miknaitis & Aksel Hallin. "Vertex Shift Studies Using ^{16}N in Salt." (2003). Technical report, Sudbury Neutrino Observatory.

[71] Kathryn Miknaitis. "Finding the AV Using neutrino data and MC." (2003). Technical report, Sudbury Neutrino Observatory.

[72] Joshua R. Klein. "Tuning the High Level Cuts for LETA." (2006). Technical report, University of Texas.

[73] SNO Collaboration. "CC Unidoc." (2001). Technical report, Sudbury Neutrino Observatory.

[74] Monica Dunford et al. "Updates to Data Cleaning for the Salt Phase." (2004). Technical report, Sudbury Neutrino Observatory.

[75] SNO Collaboration. "NC Unidoc." (2002). Technical report, Sudbury Neutrino Observatory.

[76] Joshua R. Klein. "High Level Cut Re-tuning." (2007). Meeting Presentation, University of Texas.

[77] Gordon Allan McGregor. "The Measurement of the Neutral Current Interaction at the Sudbury Neutrino Observatory." (2002). PhD thesis, University of Oxford.

[78] Stanley Seibert (2007). *Private Communication.*

[79] Pierre-Luc Drouin. "Angular Resolution Systematic Uncertainties for LETA." (2008). Technical report, Sudbury Neutrino Observatory.

[80] Scott Oser. "A Comment on Implementing the Angular Resolution Uncertainty in SIGEX" (2008). Technical report, Sudbury Neutrino Observatory.

[81] John N. Bahcall *et al. ApJ*, **555**, 990-1012 (2001). ArXiv:astro-ph/0010346v2.

[82] C. Tunnell. *Private Communication.*

[83] Olivier Simard. "PhysInt Notes" (2008). Https://manhattan.sno.laurentian.ca: MANN-7B5445. Technical report, Sudbury Neutrino Observatory.

[84] Monica Dunford. "Salt Collection Efficiency Tuning." (2006). Internal memo, Sudbury Neutrino Observatory.