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

Jeanne R. Wilson
Jesus College, Oxford

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

Trinity Term, 2004

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

Jeanne R. Wilson
Jesus College, Oxford

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

Trinity Term 2004

Abstract

The Sudbury Neutrino Observatory (SNO) has provided positive evidence for neutrino flavour conversion, by comparing neutrino fluxes measured through charged-current (CC) and neutral-current (NC) interactions in heavy water. In this thesis, data from the second phase of the experiment (the salt-phase), in which 2 tonnes of NaCl served to enhance sensitivity to the NC interaction, have been analysed. The measured ratio of CC to NC fluxes is:

$$\frac{\Phi(\mathrm{CC})}{\Phi(\mathrm{NC})} = 0.310 \pm 0.021(\mathrm{stat.}) \pm 0.024(\mathrm{sys.}),$$

which was obtained from an energy-unconstrained likelihood fit above an effective kinetic energy threshold of 5.5 MeV. This ratio measures the fraction of electron-type neutrinos with respect to the total flux of active neutrino types from the Sun and is consistent with matter enhanced (MSW) oscillations with a large mixing angle (LMA). This result is in good agreement with previous measurements by SNO and is also confirmed by the KamLAND reactor experiment.

MSW oscillations are also expected to distort the energy spectrum of electron neutrinos from the Sun and were tested for by measuring the CC energy spectrum. No significant evidence for modification of the spectrum shape was found, though the expected distortion for the LMA mixing scenario is small. Separate day and night spectra were also measured to test for matter effects in the Earth that affect electron-neutrino fluxes and spectra measured at night. For the LMA scenario, such effects are expected to be small and no significant asymmetry was detected.

Different techniques have been investigated to improve the sensitivity of SNO measurements. In particular, the technique of varying systematic parameters in the likelihood fit was found to reduce the magnitude of uncertainties on the CC and NC fluxes by 20–25%. A method to test against a range of MSW model predictions using likelihood ratios, has also been developed and tested on the salt data set.

For all my friends, who have been fantastic.

Acknowledgments

First of all I would like to thank my supervisor, Steve Biller for his endless enthusiasm and helpful advice over the past four years. I'm also extremely grateful for his efforts to ensure that all, my misplaced commas, and poor grammar were accounted for. Huge thanks are also in order for Jessica Dunmore and Charlotte Sims who have been fantastic office mates, colleagues, travel companions and above all great friends. Thank you to the rest of the Oxford group: Nick Jelley, Mike Bowler, Nick West, Simon Peeters, Wan Tseung and James Loach for all their help. Thank you also to Beverly Roger who kept me organised and made sure I always had a room in Sudbury. I have been very fortunate to have such a great group to work with.

Many other members of the collaboration have given me valuable help and advice along the way and have made trips to Sudbury enjoyable, especially Chris Kyba and Monica Dunford who have taught me many things, from the workings of the SNO electronics to the finer details of campfire cooking. Ryan, Eddy, Buzzy, and Alexander Keith have all taught me the joys, and repercussions, of the Canadian night-life.

I'd like to thank my colleagues at Queen's University and the Oxford workshop staff for all their help in the construction and implementation of the masked laserball source. Thank you also to the former Oxford graduate students; Neil McCauley has assisted me with many useful suggestions and Gordon McGregor wrote the basic MXF processor. I am grateful to many other colleagues who have helped me with the details of signal extraction, especially Bill Heintzelman and Scott Oser. In the last few months, Kevin Graham, Josh Klein and many other members of the NSP analysis groups have patiently directed me with the spectrum analysis.

I'd also like to mention everybody on the Oxford physics "pub-list". In particular, Jonas, Paul, Ankush and Steve, who have often helped me to think through various analysis and thesis writing dilemmas over a cup of tea.

Thank you to my friends in Oxford who have made life here so enjoyable. Mark Bass taught me many things you can do with LaTeX and made last-orders a vital incentive when we were writing-up. My flat-mates Emma and Harry have provided a sympathetic ear on many occasions, as have Edward, Vicki, Dan, Paul, Tom, Izzy and Nicole.

Finally, huge thanks to my family. I love you all and am proud that you have read this far - keep going!

Foreword

As SNO is a large experiment, much of the data analysis is performed in a collaborative manner. This thesis details the specific tasks performed by the author, but also draws on analyses performed by other members of the collaboration.

The main contributions for which the author was solely responsible are the measurement of Rayleigh scattering in the D_2O, described in chapter 4, and the development work on, and subsequent implementation of, the MXF maximum likelihood fitting code which is introduced in chapter 5. This code was used to perform a detailed investigation of signal extraction biases and to extract the charged-current energy spectrum from the salt data set (chapter 6). The extensions to this code to treat systematic errors in a correlated manner (chapter 7) and to perform the fit in terms of electron-neutrino energy were also developed and tested by the author (chapter 8).

The systematic studies into energy-dependent effects of energy non-linearity (appendix G), fiducial volume uncertainty (appendix H) and aborted simulation events (appendix F) were a critical requirement for the spectrum analysis that were performed by the author.

The code written to aid the task of selecting quality neutrino data for analysis was designed by the author and developed in collaboration with other members of the run selection committee.

Contents

List of Figures

List of Tables

Glossary

This glossary defines a number of acronyms, and terms frequently used in this thesis.

AV **A**crylic **V**essel.

AMB **A**nalogue **M**easurement **B**oard.

^{8}B neutrino A neutrino produced by the decay of ^{8}B to ^{8}Be in the Sun.

CC **C**harged-**C**urrent (ν interaction).

CL **C**onfidence **L**imit.

Contamination The fraction of instrumental background events in the neutrino data set.

Concentrator Reflective structure mounted around each PMT to maximise the solid angle detector coverage.

ECA **E**lectronics **CA**libration.

ES **E**lastic **S**cattering (ν interaction).

EN **E**xternal **N**eutron.

ESUM Trigger obtained by an analog sum of the signals from each PMT.

FPS **F**irst **P**ass run **S**election.

GTID **G**lobal **T**rigger **ID**entification number.

hep neutrino A neutrino produced by the capture of a proton on ^{3}He in the Sun.

NC **N**eutral-**C**urrent (ν interaction).

Laserball The main optical calibration source for the SNO detector.

^{8}Li source A calibration source that produces β^-s from the decay of ^{8}Li.

Livetime The length of a data set in days, corrected for data removed by data-cleaning criteria.

LMA **L**arge **M**ixing **A**ngle, a scenario for solar neutrino mixing.

MC **M**onte **C**arlo

MINUIT A standard minimisation package from the CERN software library.

MNS matrix The **M**aki-**N**akagawa-**S**akata unitarity matrix that described mixing between 3 neutrino flavours.

MSW effect A matter enhancement effect for neutrino oscillations named after **M**ikheyev, **S**mirnov and **W**olfenstein who developed the theory.

MXF A maximum likelihood signal extraction package developed by the author.

MTC **M**aster **T**rigger **C**ard (Digital or Analogue).

^{16}N source A source of 6.13 MeV γs produced from the decay of ^{16}N used for calibration.

NCD **N**eutral **C**urrent **D**etector.

OWL **O**ut**W**ard **L**ooking PMT.

PCA **P**MT **C**alibration.

PDF **P**robability **D**ensity **F**unction.

PMT **P**hoto-**M**ulitplier **T**ube.

PSUP **P**MT **SUP**port structure.

Pt source **P**roton **t**ritium source of high energy γs used for calibration.

PulseGT **P**ulsed **G**lobal **T**rigger.

Run Unit of collected data no longer than 24 hours in length.

RMS **R**oot **M**ean **S**quare.

Sacrifice The fraction of true neutrino events mistaken as instrumental background events and excluded from the data sample.

SNO **S**udbury **N**eutrino **O**bservatory

SNOMAN **SNO** **M**onte Carlo and **AN**alysis software.

SSM **S**tandard **S**olar **M**odel.

SPS **S**econd **P**ass run **S**election.

The following are parameters used to describe events in SNO, which are frequently used in the text.

$\boldsymbol{\beta_{14}}$ A measure of event isotopy.

$\boldsymbol{\cos\theta_{\odot}}$ The cosine of the angle between the reconstructed direction of an event, and the direction to the Sun.

ITR The in-time-ratio.

$\boldsymbol{Nhits}$ The number of PMTs signalling a hit in an event.

$\boldsymbol{R^3}$ A volume weighted measure of reconstructed event position, defined in terms of the radius of the acrylic vessel of the SNO detector.

$\boldsymbol{T_{eff}}$ The effective electron kinetic energy of an event.

Chapter 1

Solar Neutrino Theory

In 1930 Wolfgang Pauli postulated the existence of a weakly interacting neutral particle to explain the continuous β decay spectrum [1], sparking the birth of the field of Neutrino Physics. The neutrino was first detected in 1956 in an experiment by Cowan and Reines [2] through inverse β-decay of the proton by neutrinos from a nuclear reactor.

Since then neutrinos from many sources have been observed, including those created by cosmic ray interactions in the atmosphere and those created by nuclear fusion reactions in the Sun. The first indications of unusual neutrino phenomena came from the pioneering experiment of Davis *et al.* [3] which detected a smaller flux of solar neutrinos than expected. These measurements initiated the *Solar Neutrino Problem*, and were supported by results from a number of other solar neutrino experiments over the next 30 years.

The Sudbury Neutrino Observatory (SNO) is a second generation solar neutrino experiment designed to solve the solar neutrino problem, specifically probing the idea of neutrino flavour change which was first proposed by Pontecorvo and Gribov [4, 5]. Research at SNO and other second generation neutrino experiments has now unearthed neutrino properties beyond the framework of the Standard Model of Particle Physics. Indeed, a fuller understanding of the fundamental properties of neutrinos may prove vital in developing theories which go beyond this model.

In this chapter the neutrino is introduced, and the details of neutrino oscillations including recent evidence for them from SNO and other experiments are discussed.

1.1 Neutrinos in the Standard Model

The Standard Model provides a consistent theory of the fundamental particles and their electromagnetic, weak and strong interactions (for a general description see, for example, Ref. [6]). In this model, massless neutrinos exist only in left-handed doublets, partnering the charged leptons. Masses for quarks and charged leptons are generated in the Standard

Model through the Higgs Mechanism, which requires right-handed fields. With no right-handed neutrino fields, the neutrinos remain massless and their helicity state is fixed: left-handed for neutrinos and right-handed for anti-neutrinos. The latest fit to measurements of the Z^0 Boson width by the LEP experiments results in 2.9841 ± 0.0083 flavours of neutrino with mass below 45 GeV [7].

Neutrinos interact only through the Weak interaction, having neither colour nor electric charge. Reactions can proceed either through the $W^{\pm}$, converting a neutrino to its charged lepton partner, or though the neutral Z^0 boson. The large mass of these bosons makes the Weak force significantly weaker than the other forces* rendering neutrinos very difficult to detect. All three flavours of neutrino (ν_e, ν_μ and ν_τ) have now been observed through their interactions with other matter.

The Standard Model has proved extremely successful and has been confirmed by many rigorous experimental studies that test the theoretical predictions. It is generally deemed to be an incomplete theory however, as it does not incorporate the fundamental force of gravity. A number of the model input parameters cannot be explained either, for example the masses of the fundamental fermions and the fact that there are three generations. Much work has been carried out developing Grand Unified Theories (GUTs) which attempt to explain these facts and unify all four fundamental forces. Experiments probing phenomena that are not predicted by the Standard Model can provide vital inputs to these new theories.

1.2 Neutrino Mass Measurements

One phenomenon, which is not predicted in the Standard Model of Particle Physics is the finite mass of a neutrino. A number of approaches can be used to probe neutrino mass as described in this section.

1.2.1 β Decay

The lowest direct upper limits on the electron neutrino mass, m_{ν_e}, are obtainable from investigations of tritium β decay. A finite value of m_{ν_e} reduces the available phase space for the β decay and changes the shape of the emitted β spectrum. Experiments investigate this effect through measurement of the end point of tritium β decay. The existing limit on m_{ν_e} from such experiments [8, 9] is

$$m_{\nu_e} < 2.2\,\text{eV} \ (95\ \%\ \text{C.L.}).$$

*Excluding gravity, which is not incorporated in the Standard Model.

Limits on the muon neutrino mass, m_{ν_μ}, come from measurements of the two-body decay of the pion providing the limit [10]:

$$m_{\nu_\mu} < 190\,\text{keV (90\% C.L.)}.$$

The best limit on the tau neutrino mass, m_{ν_τ}, obtained from tau decay measurements is [11]

$$m_{\nu_\tau} < 18.2\text{ MeV (90\% C.L.)}.$$

1.2.2 Neutrino-less Double β Decay

It is possible to change reference frames such that a massive particle changes helicity state, therefore, massive neutrinos could have right-handed helicity. Unlike the other fundamental particles, there is no other quantum number to describe this state, such as electric charge or colour so the interpretation is ambiguous: either the right-handed helicity state is the anti-neutrino and the neutrino is its own antiparticle (a *Majorana* particle), or the neutrino is a *Dirac* particle like all other fundamental fermions, with masses generated through the Standard Higgs mechanism.

It should be noted that tritium decay experiments constrain m_{ν_e} to be at least five orders of magnitude below the electron mass, m_e, which is not explained by this mechanism. Many GUT theories predict a finite neutrino mass through the see-saw mechanism [12], which accounts for very small left-handed neutrino masses through a suitably large Majorana mass for the right-handed neutrino.

One of the most promising ways to probe the nature of neutrinos is to search for the lepton number violating process of neutrino-less double beta decay, which can only occur for Majorana neutrinos. The null results of these searches provide limits on the Majorana neutrino mass, with the most competitive limit being that from ^{76}Ge [13]:

$$m_{\nu_e} < 0.35\,\text{eV}.$$

1.2.3 Cosmological Limits

Cosmological data can also provide information on the total mass of the light active[†] neutrinos. Mass limits are derived from the way non-zero neutrino masses suppress the amplitude of density fluctuations and effect the shape of the matter power spectrum. Recent measurements from the Wilkinson Microwave Anisotropy Probe (WMAP), combined with data from the 2dF Galaxy Redshift Survey (2dFGRS), limit the neutrino mass m_ν to be $< 0.23\,\text{eV}$ assuming three degenerate neutrino species [14].[‡]

[†]Active neutrinos are those that participate in the weak interaction: ν_e, ν_μ, ν_τ.

[‡]Interpretation of these data depend on priors on other cosmological constants and also the neutrino mass hierarchy (discussed in section 1.9). Most estimates are in the region 0.23–2.5eV.

1.3 Neutrino Oscillations

Whilst not a direct measure of neutrino mass, the phenomonenon of *neutrino oscillation*, where neutrinos change from one flavour state to another, requires neutrinos to have finite mass. The idea was first hypothesised by Bruno Pontecorvo in the 1960s [4].

Mixing is a phenomena already observed in the quark sector through flavour-changing charged current interactions, such as in the $K^0 - \bar{K}^0$ meson system. The eigenstates of the weak interaction are rotated with respect to the eigenstates of the strong interaction by the Cabibbo angle. Leptons do not take part in the strong interaction but there is no reason why the weak eigenstates should be identical to the mass eigenstates.

If the mass and flavour states for neutrinos are different, there is a finite chance that neutrinos can oscillate. The simplest oscillation scenario is the mixing of two flavour states in a vacuum.

1.3.1 Neutrino Oscillations in a Vacuum

The weak eigenstates, ν_e and ν_μ can be related to the mass eigenstates ν_1 and ν_2 *via* a single mixing angle, θ_{V}.

$$\begin{pmatrix} |\nu_e\rangle \\ |\nu_\mu\rangle \end{pmatrix} = \begin{pmatrix} \cos(\theta_{\mathrm{V}}) & \sin(\theta_{\mathrm{V}}) \\ -\sin(\theta_{\mathrm{V}}) & \cos(\theta_{\mathrm{V}}) \end{pmatrix} \begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \end{pmatrix} \tag{1.1}$$

The two states ν_1 and ν_2 propagate independently, but at different speeds, owing to their different respective masses, m_1 and m_2. The relative phase increases monotonically, causing a periodic modification of the interference between the two states and results in a finite chance that after travelling some distance, a neutrino initially created as ν_e will be observed as ν_μ. The vacuum oscillation length, L_{V} is

$$L_{\mathrm{V}} = \frac{2\pi}{\nu_{\mathrm{phase}}} = \frac{4\pi E}{\Delta m^2} \tag{1.2}$$

where

$$\Delta m^2 = (m_2^2 - m_1^2) \quad \mathrm{eV}^2 \tag{1.3}$$

and E is the energy of the neutrino in MeV. The survival probability for an electron neutrino in a vacuum is:

$$P(\nu_e \to \nu_e) = 1 - \sin^2(2\theta_{\mathrm{V}}) \sin^2\left(\frac{1.27\Delta m^2 L}{E}\right), \tag{1.4}$$

where L is the distance travelled in metres.

The allowed ranges for the oscillation parameters, Δm^2 and $\sin^2 2\theta$, are limited by results of neutrino oscillations experiments. A "map" of allowed parameter space can

be built up by combining the parameter ranges compatible with different experimental results. The large distance travelled by neutrinos from the Sun, which have relatively low energies, means that solar experiments are sensitive to oscillations due to small values of Δm^2. Experiments detecting higher energy neutrinos produced in the Earth's atmosphere, are sensitive to oscillations with larger values of Δm^2.

1.4 Evidence for Neutrino Oscillation

Over the past five decades many experiments have measured neutrinos from a number of different sources. The strong evidence for oscillations in solar neutrino data is discussed in detail in sections 1.5.1 and 1.7, but first, the positive oscillation signatures seen by experiments on short-baseline neutrino beams and atmospheric neutrinos will be discussed.

1.4.1 Short-Baseline

The LSND experiment [15] searched for $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillations *via* the reaction $\bar{\nu}_e + \mathrm{p} \to e^+ + \mathrm{n}$ in liquid scintillator. The $\bar{\nu}_\mu$ beam was created from a source of pions made by proton collisions on a target. An excess of events, after background subtraction, implied a neutrino oscillation probability of

$$0.264 \pm 0.067\,(\mathrm{stat.}) \pm 0.045\,(\mathrm{sys.})\,\%,$$

which can be described by the mixing parameters:

$$\sin^2 2\theta > 0.003 \quad \text{and} \quad \Delta m^2 = (0.2 - 2)\,\mathrm{eV}^2.$$

This result was checked by the KARMEN experiment, which searched for the same signature of $\bar{\nu}_e$ appearance, but did not observe a positive oscillation signal. This null result excludes most, but not all of the "phase space" for the LSND oscillations. The MiniBooNE experiment [16] currently collecting data at Fermilab, will test the LSND result by searching for $\nu_\mu \to \nu_e$ oscillations.[§]

1.4.2 Atmospheric Neutrinos

Neutrinos produced by cosmic ray interactions in the Earth's atmosphere are created in the flavour ratio $\frac{\Phi(\nu_\mu)}{\Phi(\nu_e)} \approx 2$. The Super-Kamiokande experiment (SuperK) measured fluxes of both ν_μ and ν_e-type neutrinos through their charged-current interactions on nuclei in an H_2O target:

$$\nu_l + N \to l + X \tag{1.5}$$

[§]Assuming CPT invariance.

where l represents either an electron or a muon, N is the initial nucleus and X is the final state nucleus. The flavour of the final state lepton is used to identify the flavour of the incoming neutrino. The $\Phi(\nu_\mu)$ flux was found to be strongly dependent on zenith angle, with a large deficit of neutrinos travelling vertically upward into the detector. The observed suppression of $\Phi(\nu_\mu)$, but not of $\Phi(\nu_e)$ [17] is well explained by $\nu_\mu \rightarrow \nu_\tau$ 2 flavour oscillations with the parameters

$$\sin^2 2\theta > 0.88 \quad \text{and} \quad 1.6 \times 10^{-3} < |\Delta m^2| < 4 \times 10^{-3}\,\text{eV}^2.$$

1.4.3 Reactor Neutrinos

$\nu_e \leftrightarrow \nu_\mu$ mixing is ruled out as an explanation for the observed atmospheric neutrino fluxes by the results of experiments which have measured the flux of $\bar{\nu}_e$ produced in nuclear reactors. The earlier experiments of this type were all located a short distance (less than a few km) from the reactor, and observed no suppression to the expected neutrino flux.

1.5 Solar Neutrinos

The first solar neutrino experiments, which are described in this section, were established to study the nuclear fusion process in the Sun, rather than neutrino properties. However, these experiments yielded results in conflict with theoretical predictions. Efforts to solve this conflict, which became known as the Solar Neutrino Problem, have led to huge advances in our understanding of neutrino physics.

1.5.1 First-Generation Solar Neutrino Experiments

A number of experiments endeavoured to measure the flux of solar neutrinos through either radiochemical or water Čerenkov approaches. The radiochemical experiments look for new isotopes produced in neutrino-induced reactions on chlorine (equation 1.6) or gallium (equation 1.7) during a long (1–3 month) exposure period.

$$\nu_e + {}^{37}\text{Cl} \rightarrow e^- + {}^{37}\text{Ar} \tag{1.6}$$

$$\nu_e + {}^{71}\text{Ga} \rightarrow e^- + {}^{71}\text{Ge} \tag{1.7}$$

Water Čerenkov experiments detect electrons produced in the elastic scattering reaction:

$$\nu_l + e^- \rightarrow \nu_l + e^-, \tag{1.8}$$

Experiment	Medium	Threshold (MeV)	Predicted flux	Measured Flux
Radiochemical Experiments				
			(SNU)	(SNU)
Davis	Cl	0.814	$8.5^{+1.8}_{-1.8}$	$2.56 \pm 0.16 \pm 0.16$ [3]
GALLEX & GNO	Ga	0.2332	131^{+12}_{-10}	$74.1^{+6.7}_{-6.8}$ [18]
SAGE	Ga	0.2332	131^{+12}_{-10}	$70.9^{+5.3+3.7}_{-5.3-3.2}$ [19]
Water Čerenkov Experiments				
			$(10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1})$	$(10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1})$
Kamiokande	H_2O	7.0	$5.82\,(1 \pm 0.23)$	$2.82^{+0.25}_{-0.24} \pm 0.27$ [20]

Table 1.1: A summary of the predicted [21] and measured fluxes from the first-generation solar neutrino experiments. Where two errors are given, the first is statistical and the second is systematic. The detecting medium and energy threshold is also given for each experiment.

where l can be any active lepton flavour. These experiments make a real-time measurement of the neutrino interactions, giving information on the energy and direction of the incoming neutrinos. They have a higher energy threshold than the radiochemical experiments.

The main first-generation solar neutrino experiments are summarised in table 1.1. All these experiments measure ν_es. Only the water Čerenkov experiments can detect ν_μs and ν_τs through the elastic scattering reaction (as described in section 2.4.5) with approximately one sixth of the sensitivity to ν_es. The fluxes measured by the radiochemical experiments, given in the last column of table 1.1 are expressed in SNU, a convenient unit for solar neutrino fluxes. One SNU is 10^{-36} neutrino reactions per second per target atom.

To interpret these flux measurements, an accurate prediction of the expected flux was required. These predictions are obtained from models of the Sun.

1.5.2 The Standard Solar Model

Only a brief overview of solar modelling is given here, but more detailed information can be found in Ref. [22]. As the nearest star to Earth, the Sun has been subjected to many high precision measurements. Precise measurements of the solar mass, radius, luminosity and chemical composition are available as input parameters to solar models.

The Sun is a typical "main sequence" star powered by nuclear fusion reactions converting hydrogen into helium inside the solar core:

$$4\mathrm{p} + 2e^- \rightarrow {}^4\mathrm{He} + 2\nu_e + \mathrm{E}\,(26.7\,\mathrm{MeV}). \tag{1.9}$$

Only a small fraction of the energy produced in this process is imparted as kinetic energy to the electron neutrinos, with the majority released in the form of gamma rays and positrons. In the Sun, the fusion proceeds mostly through the pp chain which is detailed in figure 1.1, though a small amount of energy is released in the CNO cycle.¶ There are five different neutrino-producing interactions in the pp chain, whose individual neutrino energy spectra are well known from nuclear physics. It is the relative contributions of each of these interactions that solar models must predict through detailed simulations of the solar physics. The solar neutrino flux is dominated by the low energy neutrinos produced in the pp step, most of which are below the energy threshold of existing solar neutrino experiments. Higher energy neutrinos are produced in the ^{7}Be, ^{8}B and *hep* steps.

Solar models attempt to predict the fluxes from the various neutrino reactions based upon known physics and measurements of the Sun. The basis of all such models is a coupled set of differential equations for a spherical body assumed to be in hydrostatic equilibrium. The initial elemental abundance is set for a star joining the main sequence and the equations are evolved from that time to the present age of the Sun. Input parameters are adjusted so that the model reproduces the observed properties of the Sun.

The neutrino fluxes resulting from each interaction, predicted by the current preferred Standard Solar Model (SSM) referred to as BP2004 [21], are given in table 1.2 and the predicted spectrum is given in figure 1.2. Water Čerenkov experiments are only sensitive to the ^{8}B and *hep* neutrinos due to an energy threshold of around 5 MeV. The latest model predicts the total ^{8}B flux to be $5.82 \times 10^{-6}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ with a 23% uncertainty that is dominated by uncertainties in solar composition. This flux is 15% higher than the ^{8}B flux predicted by the previous model of 2001 [23], due to more accurate measurement of the $^7\mathrm{Be}(p,\gamma)^8\mathrm{B}$ cross section [24].

One of the major successes of the Standard Solar Model is its ability to predict surface vibration modes, which can be measured through helioseismology. The Standard Model provides predictions for the speed of sound in the Sun, from which it is possible to predict the frequencies of vibrational modes supported in the various solar regions. Helioseismological data on "p-modes", supported between the solar surface and the base of the Sun's convective zone, and "g-modes", which penetrate into the solar core, are in agreement with model predictions to better than 0.1%. These data limit the solar density profile required to calculate resonance parameters for the MSW effect (detailed in section 1.6). They also limit the possible temperature profile and thus constrain the

¶CNO neutrinos are difficult to detect and contribute only a minute background to the SNO experiment. They have been omitted from the discussion here.

Interaction	Predicted Flux ($\times 10^{10}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$)
pp	5.94 (1.00±0.01)
pep	1.40×10^{-2} (1.00±0.02)
hep	7.88×10^{-7} (1.00±0.16)
^{7}Be	4.86×10^{-1} (1.00±0.12)
^{8}B	5.82×10^{-4} (1.00±0.23)

Table 1.2: Fluxes for each pp chain interaction in the Sun, as predicted by the BP2004 Solar Model [21]. Uncertainties are given in parentheses as a fraction of the predicted flux.

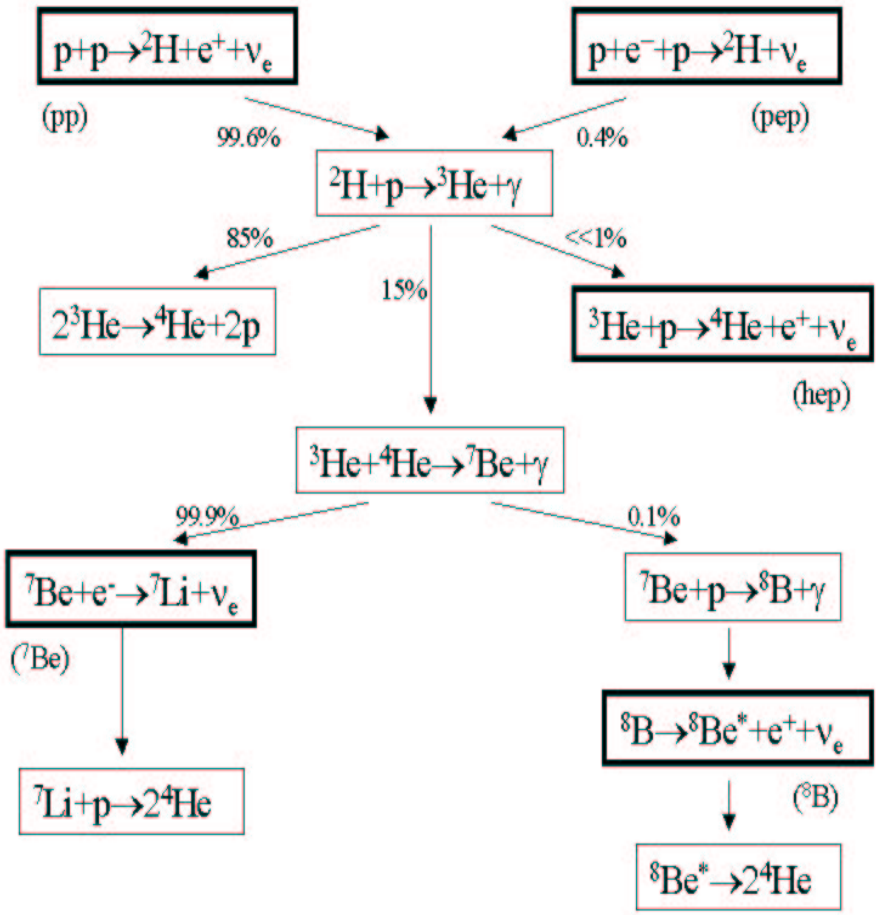


Figure 1.1: Nuclear reactions involved in the pp chain.

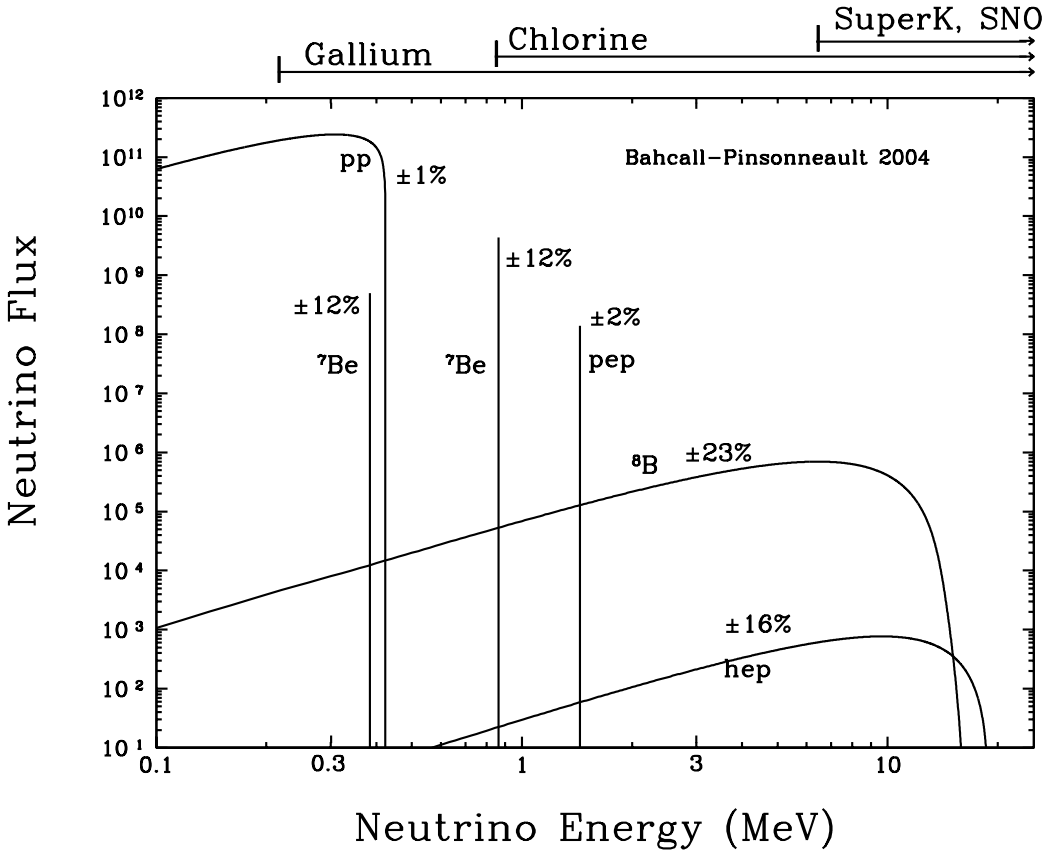


Figure 1.2: Predicted neutrino fluxes for the different solar neutrino interactions at the Earth taken from Ref. [25]. CNO cycle fluxes have been omitted for simplicity. Continuum source fluxes are given in units of neutrinos $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{MeV}^{-1}$ whilst line source fluxes are given as the number of neutrinos $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The total theoretical errors are indicated for each source.

allowed values for neutrino flux.

1.5.3 Interpretation of Results

The results of the first-generation neutrino experiments all indicate that the flux of neutrinos from the Sun is significantly lower than expected. However, the observed combination of results are not consistent with a single set of mixing parameters for vacuum oscillations.

Vacuum oscillations alone do not explain a flux suppression of > 0.5 because the $\sin^2\left(\frac{1.27\Delta m^2 L}{E}\right)$ term in equation 1.4 takes an average value of 0.5 for neutrinos created over large distances (*i.e.* a large range of L) in the Sun. This also averages over the energy dependence, so the same flux suppression would be expected for each experiment, despite the different energy thresholds.

The observed results could also be explained by matter effects which cause an energy-dependent suppression of the solar neutrino flux.

1.6 Oscillations in Matter

When propagating through matter, electron neutrinos can experience additional charge-current interactions. The effect of these interactions on neutrino oscillation was first proposed by Wolfenstein [26], and the theory was later extended to address the varying density profile in the Sun by Mikheev and Smirnov [27]. The so-called MSW effect can cause a large mixing between neutrino mass states to occur without a large vacuum mixing angle, but with an energy dependence that could explain the different observations of the various solar neutrino experiments.

As neutrinos travel through matter they can experience coherent forward scattering off electrons. For ν_e the scattering can occur either through the W (charged-current) or Z boson (neutral-current), whereas for ν_μ or ν_τ (referred to together as ν_x) only the neutral-current channel is available. This results in an effective potential that increases the effective mass for ν_e:

$$V = \frac{V_e - V_x}{2} = 2\sqrt{2} G_{\mathrm{F}} N_e, \tag{1.10}$$

in which G_{F} is the Fermi coupling constant and N_e is the electron density. The above term adds to the off-diagonal terms of the Hamiltonian which govern the propagation of vacuum mass eigenstates through matter. Mixing in matter is determined with respect to the matter eigenstates, $\nu_{1\mathrm{m}}$ and $\nu_{2\mathrm{m}}$, which depend on matter density and neutrino energy. If N_e were constant throughout the Sun, constant mixing would be observed with an oscillatory nature, as for the vacuum oscillation. However, the oscillation parameters would be modified and would depend on the neutrino energy leading to a characteristic

modification of the ν_e energy spectra. The new mixing angle, $\theta_{\rm M}$, is related to the vacuum mixing angle by

$$\tan(2\theta_{\rm M}) = \frac{\tan(2\theta_{\rm V})}{[1 - L_{\rm V}/L_e \sec(2\theta_{\rm V})]}, \tag{1.11}$$

where $L_{\rm V}$ is the vacuum oscillation length defined in equation 1.2 and L_e is the electron-neutrino scattering length:

$$L_e = \frac{\sqrt{2}\pi}{G_{\rm F} N_e}. \tag{1.12}$$

If $m_2 > m_1$ maximal mixing ($\theta_{\rm M} \approx \frac{\pi}{2}$) will occur when $L_{\rm V} = L_e \cos(2\theta_{\rm V})$ given by the resonance condition:

$$\Delta m^2 \cos(2\theta_{\rm V}) = 2\sqrt{2} N_e G_{\rm F} E. \tag{1.13}$$

This constant density situation could be applied to neutrinos crossing the mantle of the Earth, which can be probed by comparing day and night flux measurements.

In a non-uniform medium such as the Sun, N_e will change as the neutrino propagates[‖] and the Hamiltonian of the system depends on time. The mixing angle, $\theta_{\rm M}$, thus changes with time, and the states $\nu_{1\rm m}$ and $\nu_{2\rm m}$ are only instantaneous eigenstates, not eigenstates of propagation. The resonance condition of equation 1.13 may never be met for low energy neutrinos ($\leq 5\,$MeV). Some oscillation scenarios explained the suppression of neutrino flux at low energies through transitions $\nu_{1\rm m} \leftrightarrow \nu_{2\rm m}$, but this is disfavoured by recent data, which are compatible with oscillations with large mixing angles.

LMA Oscillations

The discussion presented here focuses on the Large Mixing Angle (LMA) region of parameter space, which is strongly favoured by current data (see section 1.7). The LMA scenario assumes an *adiabatic condition* where the change in density is small over the oscillation length and $\nu_{1\rm m} \leftrightarrow \nu_{2\rm m}$ transitions can be ignored. The composition of the instantaneous eigenstates depends on the instantaneous density, $N_e(t)$, but the admixture of eigenstates is fixed by the mixing at the production point, $\theta_{\rm M}(0)$.

Consider a ν_e created in a high density region at the center of the Sun. If its energy is greater than that satisfying equation 1.13 (E $> 5\,$MeV), then $\theta_{\rm M}(0) \approx 90^\circ$ and it will consist mainly of the higher mass state, $\nu_{2\rm m}$. As the neutrino propagates outward, N_e will decrease, causing $\theta_{\rm M}$ to decrease, and thus $\nu_{2\rm m}$ will pick up a significant admixture of

[‖] The radial profile of electron density in the Sun, $N_e(r)$, is a direct prediction of solar models since the different processes creating solar neutrinos occur at different radii within the core of the Sun. The instantaneous density, $N_e(t)$, is discussed here.

ν_μ. When the neutrino leaves the Sun, $N_e = 0$, $\theta_{\rm M} = \theta_{\rm V}$ and the neutrino has converted almost entirely to ν_μ.

1.7 Evidence for LMA Oscillations

To confirm the neutrino oscillation hypothesis, and determine the exact nature of the oscillations, new solar neutrino experiments were required. This section details results obtained from the two solar neutrino experiments built to specifically probe neutrino mixing phenomena.

1.7.1 Super Kamiokande

SuperK is a larger and improved version of the Kamiokande H_2O Čerenkov experiment (the results of which are presented in table 1.1), which has made measurements of solar neutrinos as well as atmospheric neutrinos. The 1496 days of data obtained in the first phase of SuperK between May 1996 and July 2001 result in the high precision measurement of the ^{8}B flux through an elastic scattering reaction in the H_2O target:

$$\Phi_{\rm ES}^{\rm SK}(\nu_x) = 2.32 \pm 0.03\,({\rm stat.})\,^{+0.08}_{-0.07}\,({\rm sys.}) \times 10^6\,{\rm cm}^{-2}\,{\rm s}^{-1}, \tag{1.14}$$

which is only 47% of the predicted flux [23].

1.7.2 Sudbury Neutrino Observatory

Sudbury Neutrino Observatory (SNO), which is described in significantly more detail in chapter 2, was built to specifically test the hypothesis of neutrino flavour change by measuring not only the electron-type neutrino flux (through the charged-current interaction described in section 2.4.3), but also the flux of all active neutrino flavours from the Sun (through neutral-current interactions described in section 2.4.4).

The SNO collaboration published two measurements of solar neutrino flux from the first phase of the experiment in which pure D_2O was used as a target. In July 2001 the first measurement of the charged-current interaction rate was reported above an effective kinetic energy threshold of 6.75 MeV [28] under the assumption of the standard ^{8}B spectrum shape [29]:

$$\Phi_{\rm CC}^{\rm SNO}(\nu_e) \;=\; 1.75 \pm 0.07\,({\rm stat.})\,^{+0.12}_{-0.11}\,({\rm sys.}) \pm 0.05\,({\rm theory}) \times 10^6\,{\rm cm}^{-2}\,{\rm s}^{-1} \tag{1.15}$$

This was compared with the high statistics flux measurement for elastic scattering events in SuperK, given in equation 1.14. The difference of $0.57 \pm 0.17 \times 10^6\,{\rm cm}^{-2}\,{\rm s}^{-1}$

between these fluxes was the first evidence for an active non-electron flavour component in the solar neutrino flux, at a 3.3σ confidence level.

In May 2002, the SNO collaboration published the first measurement of the neutral-current flux above an effective kinetic energy threshold of 5.0 MeV [30]:

$$\Phi_{\mathrm{NC}}^{\mathrm{SNO}}(\text{Phase I}) = 5.09\,^{+0.44}_{-0.43}\,(\text{stat.})\,^{+0.46}_{-0.43}\,(\text{sys.}) \times 10^{6}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1} \tag{1.16}$$

This value was in good agreement with the solar model prediction and resulted in a non-ν_e flux of

$$\Phi^{\mathrm{SNO}}(\nu_\mu, \nu_\tau) = 3.41 \pm 0.45\,(\text{stat.}) \pm 0.43\,(\text{sys.}) \times 10^{6}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1} \tag{1.17}$$

which differs from zero at 5.3σ significance.

These results were combined with other solar neutrino data to place constraints on the allowed parameters for MSW oscillations of solar neutrinos. MSW oscillation models between two active flavours were fit to the data and a χ^2 minimisation was performed for the mass difference, Δm^2, the mixing angle, θ, and the total ^{8}B flux. As the contribution of the *hep* flux is known to be small, this was fixed to the BP2000 model flux prediction of $9.3{\times}10^{-3}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. In figure 1.3 contours for χ^2 confidence levels relating to 90%, 95%, 99% and 99.73% probability ($\chi^2 = 4.61$, 5.99, 9.21 and 11.83 respectively) are shown.

For the left hand plot, only the latest results from the radiochemical experiments (as given in table 1.1) and the fluxes measured in separate zenith angle bins by SuperK were analysed. The combination of these data resulted in four distinct regions of allowed parameter space, with a best fit point at $\Delta m^2 = 4.47 \times 10^{-6}\,\mathrm{eV}^2$ and $\tan^2\theta = 5.62 \times 10^{-4}$. Inclusion of the SNO data resulted in a strong preference for the upper right hand region, known as the LMA mixing scenario. The best fit point for this analysis, which included separate day and night CC energy spectra, was $\Delta m^2 = 6.31 \times 10^{-5}\,\mathrm{eV}^2$, $\tan^2\theta = 0.398$.

In September 2003, the first results from the second phase of the SNO experiment (known as the "salt phase") were released. These were based on data obtained with enhanced sensitivity to the neutral-current reaction through the introduction of NaCl to the D_2O detecting medium. These data allowed a more precise measurement of the total active neutrino flux, free from assumptions of the underlying shape of the electron neutrino energy spectrum [31]:

$$\Phi_{\mathrm{NC}}^{\mathrm{SNO}}(\text{Phase II}) = 5.21 \pm 0.27\,(\text{stat.}) \pm 0.38\,(\text{sys.}) \times 10^{6}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1} \tag{1.18}$$

The ratio of CC to NC fluxes measured in this phase was:

$$\frac{\Phi_{\mathrm{CC}}^{\mathrm{SNO}}}{\Phi_{\mathrm{NC}}^{\mathrm{SNO}}}(\text{Phase II}) = 0.306 \pm 0.026\,(\text{stat.}) \pm 0.024\,(\text{sys.}) \tag{1.19}$$

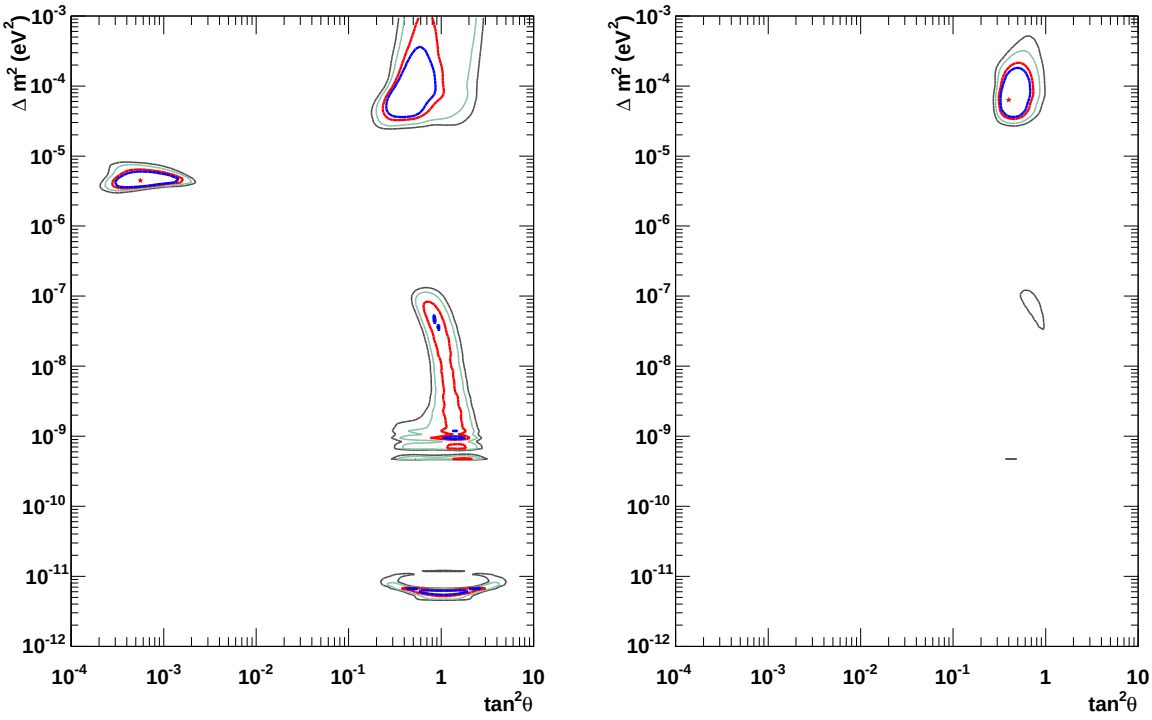


Figure 1.3: The allowed parameters for MSW oscillations of solar neutrinos from a global analysis of solar neutrino data without (left hand plot) and with (right hand plot) the SNO D_2O results. The contours represent χ^2 confidence levels relating to 90% (blue), 95% (red), 99% (green) and 99.73% (black) probability. Both figures kindly supplied by Mark Chen.

This information further reduced the allowed parameter space shown in figure 1.3. Further results from this Phase II data, including those presented in this thesis may help to further constrain the parameter space.

1.8 The KamLAND experiment

At an average reactor-detector baseline of ≈180 km, KamLAND (Kamioka Liquid-scintillator Anti-Neutrino Detector) is the first reactor experiment able to observe a suppression in reactor anti-neutrino flux. The KamLAND experiment detects anti-neutrinos from reactors spread across Japan and North Korea which produce an average flux of $1 \times 10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The flux and energy spectrum of incident neutrinos is calculated from the reactor power records of the reactors. Anti-neutrinos above an energy threshold of 1.8 MeV can be detected *via*:

$$\bar{\nu}_e + \mathrm{p} \rightarrow \mathrm{e}^+ + \mathrm{n}, \tag{1.20}$$

which gives a double coincidence signal in which the neutron is observed 220 μs after the positron through capture on another proton (see equation 1.21) in the 1000 ton liquid scintillator target.

$$\mathrm{n} + \mathrm{p} \rightarrow \mathrm{d} + \gamma. \tag{1.21}$$

The first results from KamLAND [32] for 162 ton years of reactor anti-neutrino data taken between March and October 2002, contained 54 events. Above 3.4 MeV the ratio to the number of expected events is

$$0.611 \pm 0.085\,(\mathrm{stat.}) \pm 0.04\,(\mathrm{sys.}).$$

In the context of two neutrino mixing and CPT conservation, this result excluded all oscillation solutions but the LMA solution. A maximum likelihood analysis of the observed anti-neutrino energy spectrum further reduced the allowed regions in Δm^2–$\tan^2\theta$ parameter space.

A global χ^2 analysis, including both solar data and the KamLAND results, further limited the allowed regions for oscillation parameters with a best fit point of $\Delta m^2 = 7.3 \times 10^{-5}\,\mathrm{eV}^2$ and $\tan^2\theta = 0.41$ [33]. The 1σ and 3σ allowed regions from this fit are shown in figure 1.4.

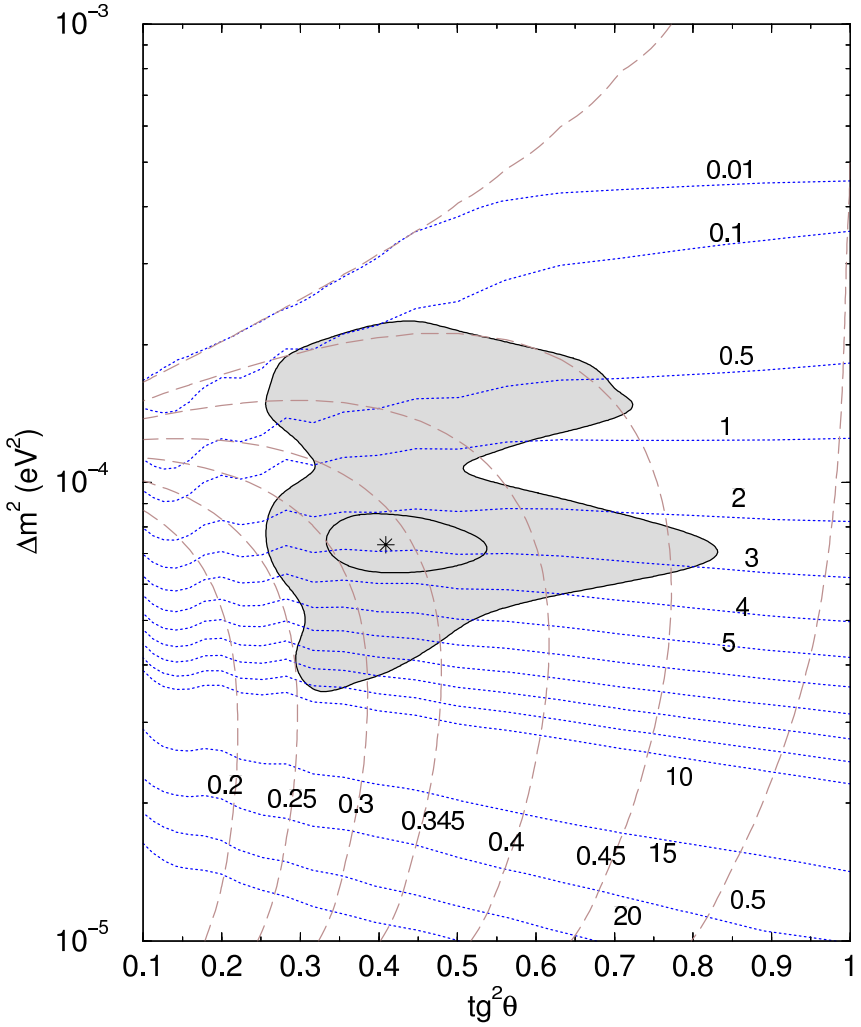


Figure 1.4: 1σ and 3σ allowed regions in Δm^2 - $\tan^2\theta$ parameter space for a combined fit to solar neutrino data and the KamLAND spectra. Predictions for the ratio of charged-current and neutral-current fluxes (dashed lines) and for the the Day-Night asymmetry (dotted lines, numbers given in %) in the SNO experiment are overlaid. Taken from [33].

1.9 The Three Neutrino Oscillation Scenario.

So far all of the neutrino oscillations discussed have been approximated to the two flavour scenario. There are, however, three active neutrino flavour states, and therefore three separate mass states (ν_1, ν_2 and ν_3), whose mixing can be described by a 3×3 unitarity matrix called the MNS (Maki–Nakagawa–Sakata) matrix:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1}, U_{e2}, U_{e3} \\ U_{\mu 1}, U_{\mu 2}, U_{\mu 3} \\ U_{\tau 1}, U_{\tau 2}, U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} = U_{\mathrm{MNS}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1.22}$$

The MNS matrix can be parameterised in terms of the three mixing angles and one charge-parity (CP) phase, δ, which, if it takes a non-zero value, would allow for the violation of CP conservation in the neutrino sector [34]:

$$\begin{aligned} U_{\mathrm{MNS}} &= U_{23} U_{13} U_{12} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13} e^{+i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned} \tag{1.23}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$.

The evidence for oscillations in different systems, which is used to place limits on the separate mixing angles and mass differences are summarised below.

- θ_{13}
  The measurement of a $\bar{\nu}_e$ flux consistent with expectations by short-baseline nuclear reactor experiments limits the mixing angle described by the second part of equation 1.23 to be small. The best limits come from the CHOOZ detector in France [35]:

$$\sin^2 2\theta_{13} < 0.10 \quad \text{for} \quad \Delta m^2_{13} = 2 \times 10^{-3}\ \mathrm{eV}^2.$$

- $|\Delta m^2_{23}|$, θ_{23}
  The terms in the first part of equation 1.23 dominate the survival probability of atmospheric ν_μ. The observed suppression of this flux can be explained by mixing between the mass states, ν_2 and ν_3.

- Δm^2_{12}, θ_{12}
  Since θ_{13} is small, the third part of equation 1.23 dominates the survival probability for ν_e, as probed by KamLAND and the solar neutrino experiments.

The sign of Δm^2_{23} is unknown, but matter effects determine the sign of Δm^2_{12} to be positive.** This results in two possible hierarchies for the neutrino mass states:

Normal hierarchy:	$m_1^2 < m_2^2 < m_3^2$
Inverted hierarchy:	$m_3^2 << m_1^2 < m_2^2$

The evidence for oscillations measured by the LSND experiment has been omitted from the above discussion. This is because these results are explained by a further Δm^2 value that cannot be accommodated in the three flavour mixing scenario. Confirmation of this result by the MiniBooNE experiment would be indicative of new neutrino mixing phenomena.

1.10 Other Possible Neutrino Phenomena

Although neutrino oscillations are the favoured explanation of the solar neutrino problem, other theories have been proposed that predict new neutrino phenomena and could produce a similar effect. Many of these theories attempt to explain the LSND result as well. It is possible that such phenomena could exist alongside neutrino oscillations, complicating the interpretation of neutrino mixing experiments. A brief description of some theories is presented here.

1.10.1 Sterile Neutrinos

Neutrino oscillations also provide the means with which to probe the existence of otherwise non-observable particles such as sterile neutrinos, which have right-handed helicity and cannot participate in the Weak interaction. It is, however, possible for sterile neutrinos to participate in neutrino mixing, resulting in a four (or more) neutrino oscillation scenario. The possibility of sterile neutrinos must be considered when interpreting neutrino oscillation experiments. Oscillation to a pure sterile state is ruled out for solar neutrinos as the total measured ^{8}B flux is in agreement with prediction, however, the possibility of oscillation to an active-sterile mixture:

$$\nu_e \rightarrow \cos(\eta)\nu_a + \sin(\eta)\nu_s, \tag{1.24}$$

is still possible in the framework of 4-neutrino mixing, where ν_a is a combination of active states, and ν_s is sterile. Global analyses place limits on the allowed active-sterile

**If $m_1 > m_2$, as defined in equation 1.11, the matter effects would serve to increase the ν_e flux with respect to the vacuum oscillation scenario for $\theta_{12} < \frac{\pi}{2}$, which would be in conflict with the solar neutrino data.

admixture [36]:

$$\sin^2 \eta \leq 0.13(0.52) \quad (1(3)\sigma \text{ C.L.}). \tag{1.25}$$

1.10.2 Spin Flavour Precession

One proposed theory requires a large neutrino magnetic moment,[††] which can interact with magnetic fields in the Sun causing the neutrino to flip into a right-handed helicity state [37, 38]:

$$\nu_{e(L)} \rightarrow \nu_{e(R)} \text{ (Dirac)}$$

This change in helicity would render Dirac neutrinos sterile to the Weak interaction. Majorana neutrinos could undergo a simultaneous flip in chirality and flavour, thus changing to a $\bar{\nu}_\mu$ or $\bar{\nu}_\tau$ that is still active to neutral-current reactions:

$$\nu_{e(L)} \rightarrow \bar{\nu}_{\mu/\tau(R)} \text{ (Majorana)}$$

This spin flavour precession (SFP) can also be resonantly enhanced by matter (resonant spin flavour precession, RSFP) resulting in an energy-dependent neutrino deficit. The observation of neutrino oscillations by KamLAND rules out RSFP as the dominant mechanism in the Sun [32].

1.10.3 Non-Standard Neutrino Interactions

Non-standard neutrino interactions are predicted by many neutrino mass models. One non-standard interaction which could possibly explain the LSND result is anomalous decay of the muon, which violates lepton number conservation by 2 units [39]:

$$\mu^+ \rightarrow \bar{\nu}_e + \bar{\nu}_l + e^+ \quad (l = e, \mu, \tau).$$

1.10.4 Neutrino Decay

Another possibility for neutrinos with non-vanishing rest mass is the decay of heavier neutrino species into a light neutrino and various additional decay products, such as a photon or a lepton-antilepton pair. As it must be mass eigenstates that decay, this is essentially a two-neutrino mixing scheme. A number of techniques may be used to search for evidence of neutrino decay, but none have given positive signals to date.

[††]Several orders of magnitude greater than the electro-weak standard model prediction.

1.11 Signatures of Matter Effects

An energy-dependent suppression of the electron neutrino flux could cause distinct signatures in solar neutrino data. Observation of the two effects described in this section would provide strong confirmation for the MSW effect.

1.11.1 Day Night Asymmetry

The MSW effect could occur in the Earth, affecting the flux of neutrinos that are observed at night. The change in density across the Earth (3–5.5 $\mathrm{g\,cm^{-3}}$ in the mantle to 10–13 $\mathrm{g\,cm^{-3}}$ in the core [40]) is not sufficient for the full MSW mechanism to occur. However, for these densities there is a range of Δm^2_{12} values for which the resonance condition of equation 1.13 is fulfilled by neutrinos in the energy region probed by Čerenkov experiments. This resonance can cause oscillations from ν_μ or ν_τ to ν_e (ν_e regeneration) increasing the flux of electron-type neutrinos at night.

Both the SuperK and SNO experiments can search for a *day-night asymmetry*, A_{DN}:

$$A_{\mathrm{DN}} = \frac{2(\Phi_{\mathrm{N}} - \Phi_{\mathrm{D}})}{\Phi_{\mathrm{N}} + \Phi_{\mathrm{D}}} \tag{1.26}$$

where Φ_{N} is the night-time flux and Φ_{D} is the day-time flux. The magnitude of A_{DN} depends on the value of Δm^2_{12}, as shown by the dotted contours on figure 1.4. For the allowed region of parameter space the expected asymmetries are small ($< 5\%$) requiring a high degree of experimental accuracy for a positive result.

The day-night asymmetry for charged-current events measured from Phase I SNO data is [41]:

$$A^{\mathrm{CC}}_{\mathrm{DN}}(\text{SNO Phase I}) = +14.0 \pm 6.3\,(\text{stat.})\,^{+1.5}_{-1.4}\,(\text{sys.})\,\%$$

If the total active neutrino flux, as measured by the neutral-current reaction is constrained to have no asymmetry, the ν_e asymmetry is found to be:

$$A^{\nu_e}_{\mathrm{DN}}(\text{SNO Phase I}) = +7.0 \pm 4.9\,(\text{stat.})\,^{+1.3}_{-1.2}\,(\text{sys.})\,\%$$

The asymmetry measured for SuperK's 1496 live-day data set is [42]:

$$A^{\mathrm{ES}}_{\mathrm{DN}}(\text{SuperK}) = +1.8 \pm 1.6\,(\text{stat.})\,^{+1.3}_{-1.2}\,(\text{sys.})\,\%$$

None of these results are significantly different from zero, but the statistical uncertainty alone on $A^{\nu_e}_{\mathrm{DN}}$(SNO Phase I) is comparable to the expected asymmetry.

1.11.2 Spectrum

Matter enhanced oscillations in the Sun can distort the energy spectrum of electron-type neutrinos. For some regions of MSW parameter space, these distortions may be visible in the energy range probed by the SNO experiment. Figure 1.5 shows how the shape of the charged-current and elastic scattering energy spectra in SNO are expected to change for an MSW effect with the mixing parameters $\Delta m^2_{12} = 7.3 \times 10^{-5}\,\mathrm{eV}^2, \tan^2\theta_{12} = 0.41$.

The predicted spectrum for the favoured LMA solution is higher than the undistorted spectrum at low energies. The distortion is less pronounced in the elastic scattering spectrum because the mapping between neutrino energy and Čerenkov electron energy is smeared out by the kinematics of this interaction. The size of the effect depends on the exact values of Δm^2_{12} and θ_{12}. Figure 1.6 shows the ratio of the normalised charged-current spectra to the non-distorted case for different points in MSW parameter space. For the mixing parameters shown, matter effects are predicted to increase the flux by less than 5% at 5.5 MeV. To test for these effects the combined statistical and systematic uncertainties in the measured spectrum must be less than 5% at low energies.

1.12 Summary

Over recent years, many experiments have contributed to the growing evidence for neutrino oscillations. The SNO experiment finally solved the *Solar Neutrino Problem* with direct evidence for neutrino flavour mixing. The KamLAND reactor experiment later provided confirmation of the Large Mixing Angle oscillation solution, which is favoured by solar data. Following these exciting revelations in neutrino physics, precision measurements are required to obtain a fuller understanding of neutrino mixing phenomena. One such measurement is that of the solar neutrino energy spectrum, which will be presented in this thesis.

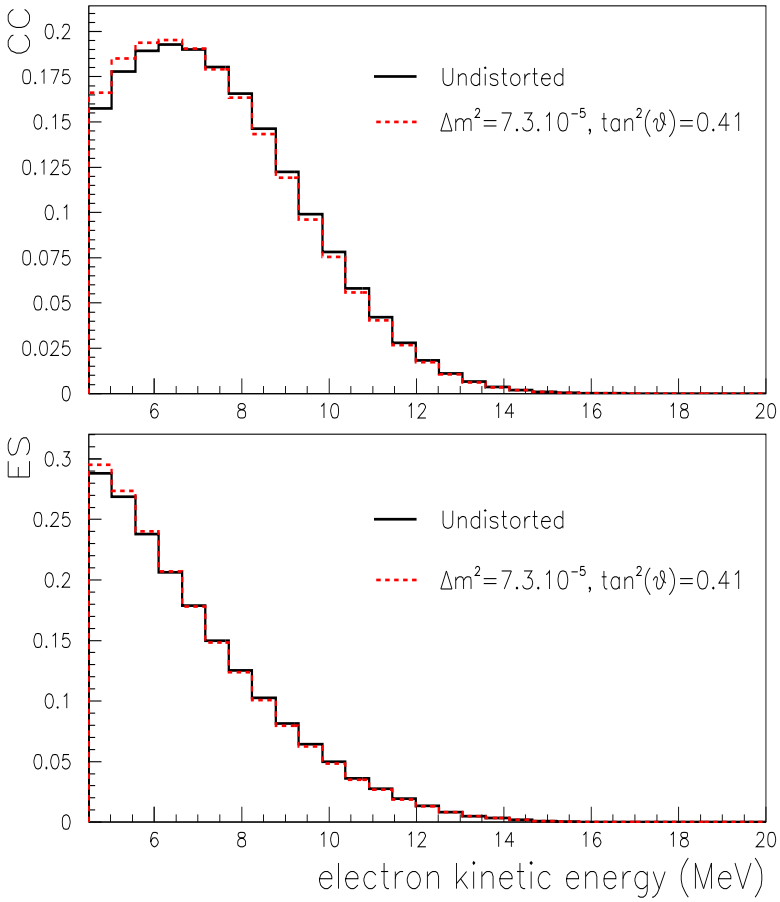


Figure 1.5: Predictions for the SNO charge-current and elastic scattering energy spectra from an undistorted ^{8}B spectrum (black solid lines), and for MSW distortions with the parameters $\Delta m^2_{12} = 7.3 \times 10^{-5}\,\mathrm{eV}^2$ and $\tan^2\theta_{12} = 0.41$ (red dashed lines).

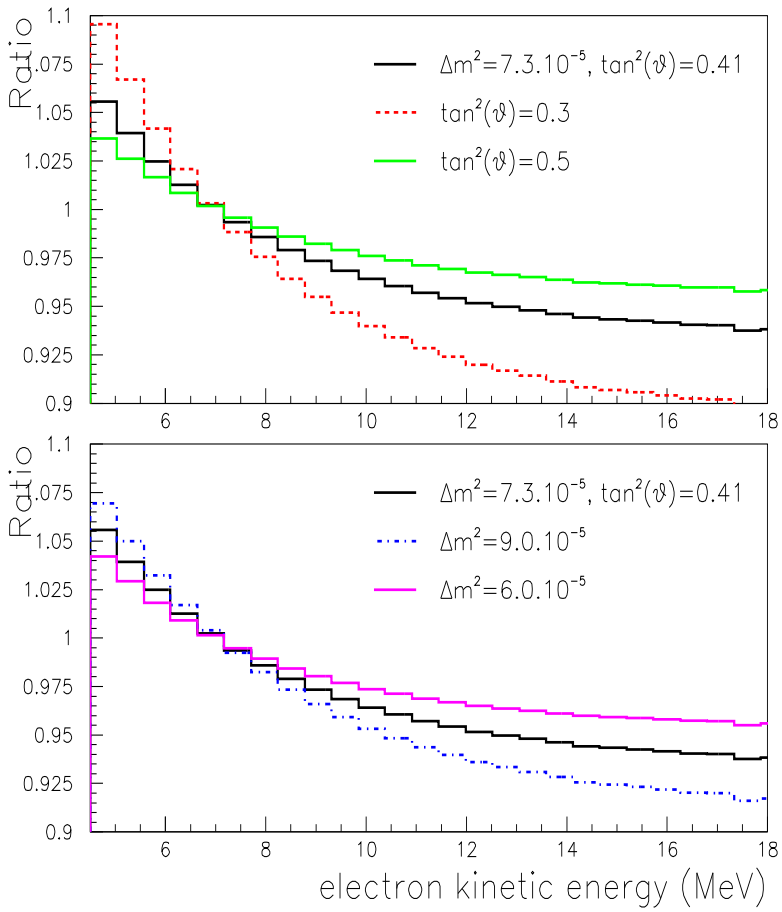


Figure 1.6: The ratio of predicted charged-current spectra, for various MSW scenarios, to the undistorted spectrum. The black solid lines give the ratio for the parameters used to create figure 1.5. The upper panel shows the effect of changing $\tan^2\theta_{12}$ on the ratio and the lower panel demonstrates the dependence of the ratio on Δm^2_{12}.

Chapter 2

The Sudbury Neutrino Observatory

The Sudbury Neutrino Observatory (SNO) is a kilo-tonne heavy water Čerenkov detector located over 2 km underground in the INCO Ltd. Creighton mine near Sudbury, Ontario. The experiment was proposed by Herb Chen [43] to solve the Solar Neutrino Problem. Although the Sun produces only electron-type neutrinos, the target of heavy water used in SNO is sensitive to all active flavours of neutrino. With this unique capability, SNO can make a direct measurement of the non-electron type neutrino flux from the Sun and directly test the hypothesis of neutrino oscillation.

2.1 The SNO Detector

The SNO detector is located at a depth of 2092 m, which is equivalent to 6010 m of water. At this depth the number of cosmic-ray muons passing through the detector is limited to about 100 per day. The detector (shown in figure 2.1) is situated in a large, barrel-shaped cavity, 22 m in diameter and 34 m in height. The 1 kilotonne target of ultra-pure D_2O is contained within a transparent, acrylic vessel (AV) 12 m in diameter and 5.5 cm thick. For analysis purposes cartesian coordinates are defined with respect to the center of the AV with a vertical z axis, such that $(x, y, z) = (0, 0, 0)$. This coordinate system will be used throughout this thesis.

At the top of the AV is a chimney (Neck) 1.5 m in diameter and 6.8 m tall which allows limited access to the D_2O volume. The AV is constructed from 125 acrylic panels made of special UV-transmitting acrylic, which were bonded together underground. The acrylic structure is supported from above by 10 vectran ropes, which attach to thicker panels around its equator. The remaining volume is filled with ultra-pure H_2O which acts as a cosmic ray veto and as a shield from naturally occurring radioactivity from both the construction materials and the surrounding rock. The rock walls have also been coated with a urylon plastic coating in order to keep the water in and radon out.

The Čerenkov signals are detected by 9456 20 cm diameter Hamamatsu R1408 photomultiplier tubes (PMTs), constructed specially for SNO with low radioactivity Shott glass photocathodes. Each PMT has 9 dynodes and is operated at a positive anode voltage of around 2 kV, achieving a nominal gain of 2×10^7. There are 9438 inward looking PMTS, 49 of which are operated as low gain channels to extend the dynamic range of the detector at high light intensities. In addition there are 91 OutWard Looking (OWL) PMTs which serve as a muon veto. These are all mounted on the PMT Support Structure (PSUP), a 17.8 m diameter, geodesic sphere supported by 16 steel cables which is located in the H_2O volume surrounding the AV.

4 PMTs have also been mounted in the Neck region of the detector to veto events caused by a static electric discharge in the gas above the D_2O. There are also 23 PMTs mounted on the Berkley Underwater Tube Test Sled (BUTTS) suspended in the H_2O outside and approximately level with the top of the PSUP. These tubes face inward and are located within easy access for studies of PMT aging. A 27 cm diameter reflective *Concentrator* is mounted on each inward looking PMT to increase the effective photocathode coverage to $\approx$ 59%. The collection efficiency of each PMT is around 15% and has a roughly constant acceptance out to an angle of 50° from the central axis. After this, the PMT angular response drops off sharply so that each PMT sees all of the D_2O volume but little of the H_2O.

Fourteen magnetic compensation coils serve to cancel the vertical component of the Earth's magnetic field. This correction increases the collection efficiency of the PMTs by around 10% on average.

2.1.1 Electronics

The SNO electronics system is responsible for generating the event trigger and recording PMT signals continuously without significant dead time. Although the actual solar neutrino event rate is very small, the system must cope with large rates from backgrounds and calibrations, or in the event of a Supernova. The detector is organised in a modular fashion with every PMT connected to a separate *Channel* in the electronics. 32 channels are grouped together onto a *Card*, and 16 cards are grouped together into a *Crate*.

A photon striking a PMT has a certain probability of producing a photo-electron, which is then accelerated down the dynode chain of the PMT. The resultant pulse is $\approx$12 ns wide and a few millivolts in amplitude. This pulse then travels down a 75 m waterproof coaxial cable to the PMT Interface Card (PMTIC). The PMTICs provide high voltage to each channel and interface directly with a Front End Card (FEC). The FEC is responsible for processing, digitising and storing the charge and timing information from each PMT.

If a PMT pulse is of sufficient amplitude, it will fire a discriminator and initiate

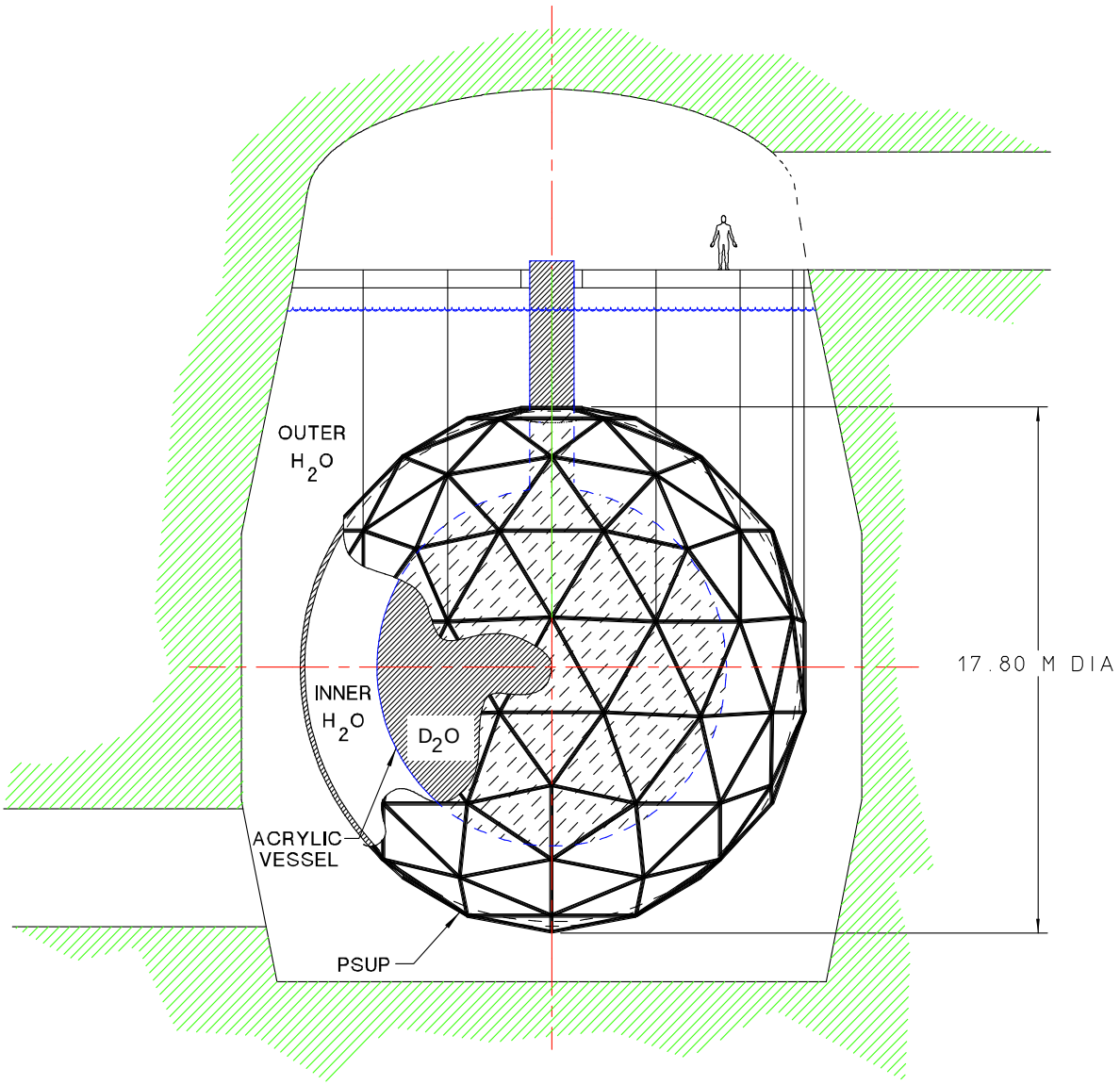


Figure 2.1: A cross-sectional view of the SNO detector

the production of two separate pulses: one integrated over a 20 ns time window (called NHIT20) and one integrated over a 100 ns time window (called NHIT100). The main physics analysis hardware trigger is the NHIT100, chosen to accommodate Čerenkov light reflections off the PSUP which occur over approximately 80 ns. The trigger pulses are summed separately over each card and crate. The Master Trigger Card Analog (MTC/A) sums the pulses from all 19 crates into a single analog pulse with a height proportional to the number of PMTs that have fired within 100 ns of each other. A third pulse, called the ESUM trigger, is obtained by an analog sum of the signals from each PMT, without the digitisation of the discriminator. SNO uses a number of different triggers,* not all of which will be detailed here. One important trigger is the "pulsed global trigger" (PulseGT) which is fired at a rate of 5 Hz to sample the nominal activity within the detector. This provides a measure of the dark-noise rate of the PMTs. There are also triggers which can be activated by external calibration sources, and triggers designed for calibration of the electronics which force readout of certain channels.

If the amplitude of any trigger pulse exceeds the threshold set on the relevant discriminator, a pulse is sent to the Master Trigger Card Digital (MTC/D). If the electronics has been programmed to readout when this trigger fires (the trigger has been "masked in"), a *Global Trigger* is sent to all the crates in the system. The timing and charge of all PMTs that fired is stored along with a 24-bit Global Trigger IDentification number (GTID) in a 16 cell analog storage array until it can be read out. Additionally, for each global trigger, the MTC/D creates an event header to store the event time, the GTID and information on which triggers fired. Two event times are stored based on a 10 MHz clock which is synchronized with a GPS receiver on surface, and also a 50 MHz clock based on a local oscillator.

Timing information for a PMT is stored on the capacitor of a Time to Amplitude Converter (TAC), which starts ramping when the discriminator for a channel is fired and stops when the global trigger is received. If no trigger is received, the TAC is reset. The pulse is fed into two separate channels of a charge integrator: one low gain, one high gain. These charge signals are then integrated over a long and short time window which cover prompt and delayed light. Only one of the low gain charges is stored depending on which time window is set in the electronics.† Both the short (QHS) and long (QHL) time window integrated charges are stored for the high gain signal. The charges are digitised by Analog to Digital Converters (ADCs) and stored along with the TAC and GTID for each event in a *PMT bundle*, which is passed along to the Data AcQuisition (DAQ) system.

*There are 3 adjustable discriminators for the NHIT100, producing 3 separate triggers: LO, MED and HI. Typically the NHIT MED threshold is set to 16–18 hits equivalent to an event energy of around 2 MeV. Likewise there are more than one NHIT20 and ESUM trigger.

†Typically the longer integration period, QLL, is used.

Element	D_2O threshold (g/g)	H_2O threshold (g/g)
^{232}Th	3.7×10^{-15}	3.7×10^{-14}
^{238}U	4.5×10^{-14}	4.5×10^{-13}

Table 2.1: Required upper thresholds for the levels of radioactive contaminants in the SNO detector

A program called the *Builder* combines all PMT bundles with the same GTID along with the relevant run header to assemble events. Occasionally, a PMT bundle will not be successfully allocated to any event. The builder stores up 80 of these unattached PMT bundles, known as *Orphans*, and writes them out together. Another system, called SHARC (SNO Hardware Acquisition and Readout Control), is used as an interface for detector control and monitoring. SHARC also stores detector settings such as channel thresholds and PMT status in *DQXX banks* which are required for data analysis.

The SNO detector can operate continuously but for organisational purposes the collected data are separated into units called *runs* which have a maximum duration of 24 hours. Runs may be divided into *sub-runs* in order to limit the size of stored data-files. All neutrino analysis is carried out at "run-level" meaning that only complete runs are analysed. The selection criteria for neutrino data are described in appendix B.

2.1.2 Water Systems

Low levels of ^{232}Th and ^{238}U can be found naturally in all materials and can contribute a background to the neutrino signals measured in SNO through their decay to particles such as βs and γs. For this reason, the entire SNO detector was built of low radioactivity materials, and stringent clean-room procedures are maintained in the underground laboratory to prevent radioactivity from mine dust and air entering the experiment. The purpose of the SNO water systems is to both maintain and measure the low levels of radioactivity required for neutrino analysis.

There are separate systems to control and measure the content of Th, U and Rn in the SNO D_2O and H_2O regions. The target levels for these elements are set so that the background they produce does not constitute more than 5% of the expected SSM signal. The required upper limits are summarised in table 2.1.

D_2O Systems

The kilotonne target of heavy water, loaned to SNO by Atomic Energy of Canada (AECL), has an isotopic purity of greater than 99.9% [44]. The D_2O systems must maintain this purity whilst keeping the levels of radioactive contaminants below the required thresholds

and maintaining the water temperature at a constant 10°C. This low temperature serves to both reduce the PMT noise rate, and the possibility of biological growth inside the water volume.[‡] Impurities are removed by periodically circulating the D_2O through a reverse osmosis system and an ultrafiltration unit.

The D_2O is assayed regularly for density, pH, conductivity, turbity, anions, cations. total organic carbon and suspended solids. Two different assay techniques are employed. In the first type of assay, Ra and Pb are stripped out of the water as it passes through a polypropylene column of acrylic beads coated in a magnesium oxide compound (MnO_x). This has an extraction efficiency of around 95% for pure D_2O and about 81% for salted D_2O for Ra [45, 46, 47]. The Ra daughters in the column are then counted offline using an electrostatic device in which charged Po ions from Rn decay are deposited on the surface of a Si alpha counter.

The second type of assay removes Th, Ra and Pb isotopes through a seeded ultrafiltration technique (SUF) in which 100 tonnes of water are passed through hydrous titanium oxide (HTiO) coated micro-filters at a rate of 150 litres per minute [48]. After exposure the columns are eluted with nitric acid, mixed with liquid scintillator and counted using β–α delayed coincidence counters. ^{224}Rn from the Th chain are seen through the coincidence of ^{212}Bi β-decay and ^{212}Po α-decay, whilst the coincidence of ^{214}Bi β-decay and ^{214}Po α-decay indicate isotopes from the U chain. The efficiency of this assay process was more strongly affected by the addition of salt to the D_2O and these measurements were thus used as a cross-check of the MnO_x assay readings for the data obtained with salted D_2O.

There is approximately 20% uncertainty in the assay radioactivity measurements, but "in-situ" measurements of radioactivity levels have also been carried out using PMT Čerenkov data. These are in good agreement and indicate that the required goals for radioactive purity have been achieved or surpassed [49].

Before entering the detector, the D_2O is de-gassed to remove radon. To prevent radon from the air in the laboratory entering the water, both the D_2O and H_2O are protected by a nitrogen cover-gas system. The remaining radon content in the water is measured using a "process de-gasser". 6 tonne samples taken from a selection of regions in the D_2O or H_2O are passed through a small vacuum de-gasser at a rate of 20 litres per minute to remove gases. The ^{222}Rn in this gas sample is frozen out in a U-shaped trap cooled with liquid nitrogen and passed to a ZnS coated scintillation cell for counting.

[‡]Without cooling the water temperature would rise to that of the surrounding rock ($\approx 41°$).

H_2O Systems

The main purpose of the light water is to provide shielding from radioactivity in the surrounding rock and PMT array, but it must also be low in radioactivity with excellent optical properties. Ultra-pure H_2O can leach elements out of solid surfaces and can also support biological activity on those surfaces. Thus the H_2O in the SNO detector is continuously recirculated and repeatedly assayed, using the same assay techniques as for the D_2O.[§]

Before entering the detector the water is filtered through a series of membranes and a reverse osmosis unit, and cooled to 10°C. It is also passed through a Ultra-Violet unit to eliminate biological contamination. O_2 and Rn are removed, along with other gases, in a process de-gasser. However, it was found that removing all gases from the H_2O resulted in low pressure in the underwater connections to the PMTs, compromising their breakdown voltages. Thus, the water is re-gassed with pure nitrogen.

2.2 Calibrations

The energy of a particle detected in SNO is proportional to the amount of light deposited in the detector, which in turn is proportional to the number of PMTs that fire. The trajectory of the particle is determined from the relative times that the PMTs fire. In order to accurately reconstruct the energy, position and direction of an event extensive calibration of the SNO detector is required. To do this, it is necessary to place a variety of sources inside the detector. Sources are deployed using a manipulator system (shown in figure 2.2), which uses a series of ropes and pulleys to move the source in an East-West and/or North-South plane. Sources can be deployed through the neck into the D_2O volume or through one of 6 "guide tubes" (2 of these are shown in figure 2.2) into the H_2O region inside the PSUP.

Table 2.2 gives a list of calibration sources used in SNO. The following section deals with the main calibration efforts in SNO, though a large amount of calibration data from many different sources is used to check detector properties and limit systematic uncertainties.

2.2.1 Electronics Calibration

There are three components to the electronics calibrations.

- **ECAL**

  This calibrates the fundamental aspects of the channel properties, adjusting the

[§]Flow rates and exposures differ for the H_2O assay measurements.

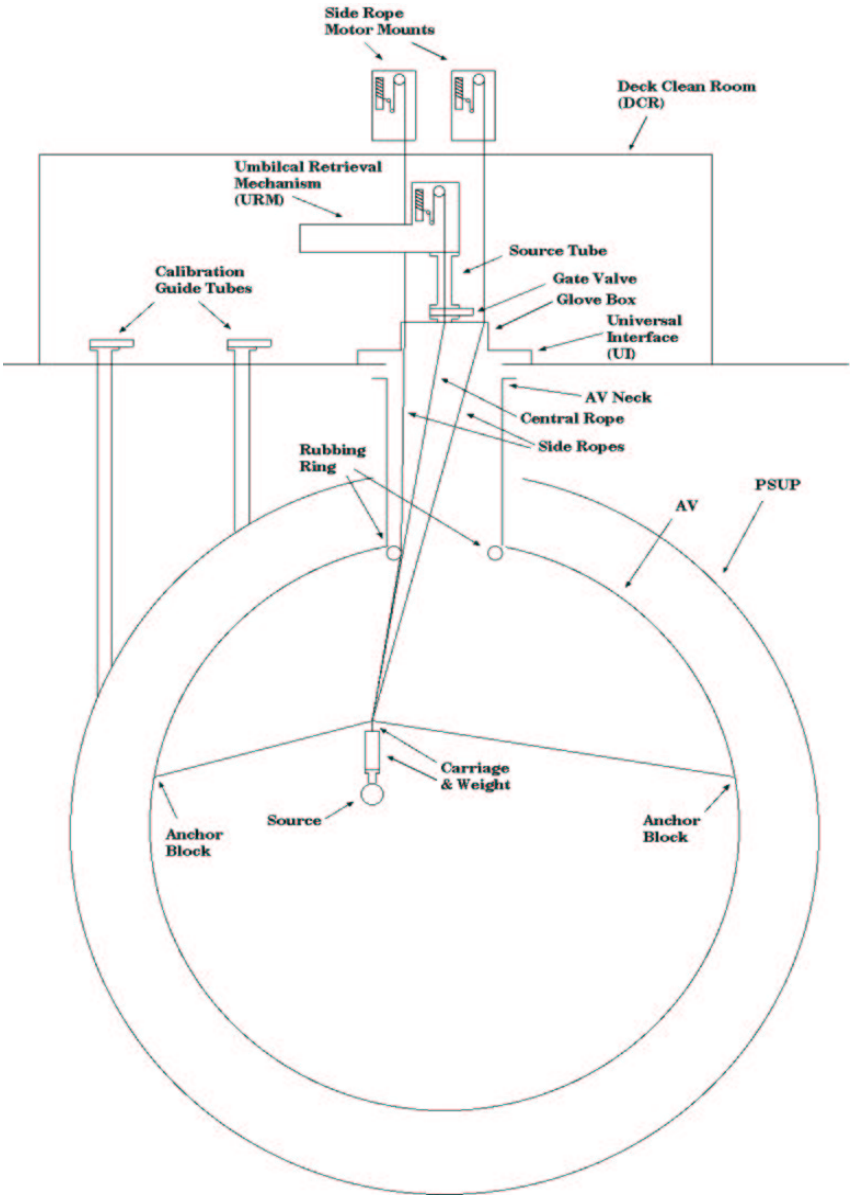


Figure 2.2: This figure shows the calibration source manipulator system and also the guide tubes that allow access to the detector for calibration purposes.

Source	Specifications	Purpose
Laserball	Multi-wavelength, isotropic pulsed light source.	PCA and optical calibrations.
^{16}N [50]	Triggered ^{16}N 6.13 MeV γ-ray source.	Primary energy calibration source.
Acrylic neutron source	Un-triggered ^{252}Cf neutron source encapsulated in acrylic.	Calibrate the detector response to neutrons.
Acrylic U and Th source	Un-triggered U and Th sources encapsulated in acrylic.	Calibrate detectors response to low energy backgrounds.
^{8}Li source [51]	Triggered source of ^{8}Li giving electrons with a spectrum similar to ^{8}B (end-point $\approx$14 MeV).	Confirmation of energy response and event reconstruction.
Pt source [52]	p+^{3}H $\rightarrow^{4}$He + 19.8 MeVγ	High energy calibration.
AmBe source	AmBe source producing 4.14 MeV γ and n in coincidence.	Low energy and neutron calibration.
Radon spike	Controlled injections of Rn into either the D_2O or H_2O.	Low energy background calibration.

Table 2.2: Calibration sources used in SNO.

channel timing and discriminator thresholds using special circuitry and software built into the system.

- **ECA** - Electronics Calibration
  The charge pedestals (the "zero charge" readings from the ADCs) and TAC slopes (rise-time of the pulse produced by the time to amplitude converter) are calibrated twice a week, as these tend to change with time. To do this the discriminators are forced to fire. After a fixed delay the global trigger is also fired. The amplitude measured for this precise time interval gives an accurate calibration for the TACs.

- **PCA** - PMT Calibration
  The PCA measures the timing response for each PMT. The size of a charge pulse created by a photo-electron can vary considerably from one PMT to another. As the discriminators fire at a fixed voltage, larger pulses will fire the discriminators earlier (referred to as "discriminator walk"). This difference in firing times between PMTs can be as large as 2 ns and is overcome by applying a correction to the PMT time based on the size of the charge pulse measured. The correction is determined using data from the *Laserball* source which provides isotropic light in narrow pulses compared to the PMT timing distribution.

2.2.2 Optical Calibration

Optical calibration involves the measurement of the properties of the D_2O, H_2O, AV and PMTs. These include the attenuation in the water and acrylic as a function of wavelength, and the relative sensitivity of the PMTs to the direction of incident light. The *laserball* is deployed at many different locations in the detector using a number of different wavelengths. A global optics model is fitted to these data and the optical properties are extracted. As detector optical properties are known to change with time, optics scans are repeated at regular intervals. This is discussed in more detail in chapter 4.

2.2.3 Energy Calibration

The amount of light produced by the laserball is not constant and, therefore, this source is unsuitable for calibrating the quantum efficiency of the PMTs (the probability that a tube produces a hit signal when struck by a photon). Instead, the ^{16}N source, which produces 6.13 MeV γ-rays is used. The γs undergo Compton Scattering in the detector to produce detectable energetic electrons. These energy response calibrations are also repeated on a regular basis. The Pt source provides a high energy calibration point to

test the energy calibration at high energies,[¶] whilst the ^{8}Li source is also used to check the energy calibration across the region of interest. Chapter 3 describes an analytical calculation carried out by the author to verify this extrapolation.

2.3 Software

Interactions in the SNO detector are simulated in the Fortran based SNO Monte Carlo and ANalysis software (SNOMAN), which is also the collaboration's principle analysis package. SNOMAN contains a number of separate "processors" which communicate through a central data structure managed by the CERNLIB package ZEBRA [53].

2.3.1 Analysis Tools

Data is written to tape in the form of packed ZEBRA banks. The analysis procedure starts with SNOMAN processors that unpack this data and calibrate it using stored electronic calibration constants to produce PMT hits with corrected times and charges. A similar set of processors prepare simulated data but, after the calibration processor, analysis procedures are identical for data and Monte Carlo.

There are a number of different "fitter" processors which derive the event position and direction using the relative timing and angular information of the PMT hits. Unless otherwise stated all reconstructed event positions (given in terms of (x, y, z) or R) referred to in this thesis were obtained using the QPDF fitter [54].

A number of other processors exist to obtain additional information from the PMT hits, for example the *RSP* processor which is described in more detail in section 2.4.2.

2.3.2 Monte Carlo Particle Propagation.

In SNOMAN the simulation of the propagation of electrons and gammas is handled by EGS4 (electron gamma shower) code [55], which includes effects of

- Collision energy loss of electrons
- Bremsstrahlung radiation
- Multiple Coulomb scattering from atomic nuclei using Molière theory
- Bhabha and Møller scattering
- Coherent Rayleigh scattering

[¶]Unfortunately it was not possible to operate the Pt source in the Salt phase due to the amount of neutrons also produced by this source.

- Compton Scattering
- Pair production
- The photo-electric effect
- Positron annihilation

Correct modelling of an electron's scattering is vital to extract the correct angular distribution of Čerenkov photons. Electrons undergo multiple Coulomb scattering as they propagate, which EGS4 simulates using Molière theory. The electron path is broken into steps, each associated with a fractional change in the kinetic energy.

EGS4 treats energy loss in a continuous manner below a given cut-off point. This cut-off is an input parameter of the code. Any interaction giving a secondary particle (δ-ray or γ) of energy above this cut-off is considered discrete and these particles are then propagated as part of the shower. Otherwise, the interaction just contributes to a continuous energy loss and is assumed not to affect the direction of propagation.

If each track segment is straight, then all the Čerenkov photons emitted from that segment will be in the same cone, defined by the Čerenkov angle. Thus, it is apparent that the number of photons depends on the number of cones, which is directly related to the step size. To remove this artefact, the local direction cosines of the track are linearly interpolated between successive tracks giving a curved track segment instead of a straight one. Using this interpolation, a step size of 0.01 was found to be applicable [56].

Analysis of the salt data makes use of event topology and was found to be extremely sensitive to the angular distribution of Čerenkov photons. An initial discrepancy between calibration data and Monte Carlo simulation was found to be due to the omission of Mott terms in the electron scattering cross section in EGS4. The angular distribution of electron scattering in the Monte Carlo was corrected to first order to account for this [57].

2.3.3 SNOMAN Generation of Čerenkov Light

The Čerenkov radiation from each segment of an electron path is determined from the asymptotic formula for light yield [58, 56] which can be divided by photon energy, $\hbar\omega$, to give the equivalent photon yield:

$$\frac{dI}{d\omega} = \frac{\omega e^2 L \sin^2\theta_c}{c^2} \tag{2.1}$$

where I, is given per unit angular frequency, ω, and is proportional to path length, L. This formula is valid for a straight section of track using classical radiation theory. The number of photons emitted is then sampled from a Poissonian distribution and the creation

points of these photons are positioned randomly along the track segment. All photons are emitted at an angle θ_c to the electron track direction, which is interpolated as described in section 2.3.2.

2.3.4 Neutron Transport

Neutron transport code is required to calculate the mean free path of a neutron in a medium and to simulate any interactions that might occur. SNOMAN makes use of the MCNP neutron transport code, developed at Los Alamos for this purpose [59]. The main elements of the MCNP code were incorporated into SNOMAN to transport γs, $e^{\pm}$s, $\mu^{\pm}$s, and neutrons. Other particles generated by neutron collisions such as αs are not propagated, but interaction vertices indicate which particles were generated.

Data tables are used to provide information on partial and total interaction cross sections as a function of neutron energy, the energy-angle spectrum of the emergent neutrons, and other interaction data. Details of the verification of SNOMAN neutron transport code through comparison with MCPN can be found in [60].

2.3.5 Monte Carlo Data Acquisition

The purpose of the Monte Carlo Data Acquisition (MCDAQ) is to extract the data that would be available for a real event from the vast amount of information in the ZEBRA banks after event simulation. The idea is that when processors act on the new banks, they cannot distinguish between data and Monte Carlo. MCDAQ calls routines to simulate the electronics effects including:

- PMT noise - A routine separates generated hits in the event into groups separated by more than the width of the trigger window. Noise hits are then added on random PMTs at random times distributed between an input offset before the first hit in the group and a different offset after the last hit in the group.

- Multiple hits - PMT hits on the same tube are combined into single charge (Q) measurements if the hits occur within the discriminator pulse width.

- Triggering - A routine looks through a time ordered list of the PMT times after walk effects have been added. The trigger generation is a simulation of the analog sum actually used in the detector. The code then steps through time in 0.25 ns increments, adding up the amplitudes of all the trigger pulses currently active. If the summed signal exceeds the hardware threshold, it creates a global trigger which is latched by the 50 MHz clock. If no triggers are found, the event is discarded.

- Discriminator walk - To replicate this effect, a charge-correlated delay is applied to the simulated firing time of each PMT when simulating the trigger.

2.3.6 Database

MCDAQ also makes use of calibration constants and parameters describing detector conditions stored in the database, SNODB, which is based on the CERNLIB package HEPDB [61]. SNODB contains zebra banks called *titles* that encode information about detector status and configuration, and information derived from calibrations required to accurately interpret the SNO data. Each bank has an effective time validity to ensure that the relevant constants are applied to the data.

2.4 Interactions in SNO

In this section the three different types of neutrino-interaction observed in SNO are discussed along with other types of event that act as backgrounds to these neutrino signals. The majority of interactions[‖] are observed through Čerenkov light.

2.4.1 Čerenkov Light

When a charged particle has a velocity greater than the local phase velocity of light in the medium through which it is travelling, Čerenkov light will be produced. The charged particle detected from neutrino interactions is an electron. The mean number of photons, N, produced per unit path length, dx, per unit energy, dE, for an electron with velocity $v = \beta c$ is given by:

$$\frac{d^2N}{dEdx} = \frac{2\pi\alpha}{\lambda^2}\left(1 - \frac{1}{(\beta n(\lambda))^2}\right) \tag{2.2}$$

where λ is the wavelength of photons emitted, α is the fine structure constant and n is the refractive index of the medium, which is a function of wavelength and hence, photon energy.

Čerenkov light is emitted in a cone around the charged particle direction with a half angle given by:

$$\theta_c = \cos^{-1}\left(\frac{1}{n\beta}\right) \tag{2.3}$$

[‖]With the exception of neutron interactions in the Neutral Current Detectors used in phase III of the experiment.

for D_2O $n \approx 1.34$, which results in a minimum electron energy threshold of 0.767 MeV though, in practice, analysis thresholds are set much higher due to low energy backgrounds from radioactive decays.

2.4.2 Event Energy

A number different variables that describe event energy are used in this thesis and are defined in this section.

The number of PMTs hit in the 100 ns time window defining an event, *Nhits*, is directly related to the energy deposited in the SNO detector. However, this observable is correlated to a number of other factors, such as the vertex position and direction of the event. The quantity is also sensitive to the number and location of PMTs that have been disabled during a given run, and the dark noise rate of the PMTs.

The RSP processor [54] was developed to provide improved mapping to energy. It calculates an equivalent electron total energy in MeV for each event using the number of "in-time" hits, arriving within ± 10 ns of the prompt time of the event.** The number of hits due to dark current in the PMTs is subtracted from this number, based on predictions using measurements from the PulseGT. The equivalent number of hits for an event at the center of the detector, N_{eff}, is then calculated by accounting for the PMT optical response, the number of PMTs online and the attenuation of photons in the various detector media. A mapping from N_{eff} to effective energy, E_{RSP} is then carried out using Monte Carlo simulations of electrons at the center of the detector. The mapped energy is convolved with a response function making E_{RSP} a more continuous variable than *Nhits*, which is an integer.

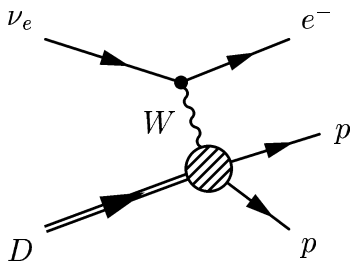
For the data set discussed in this thesis, a correction was required in order to give a good agreement between E_{RSP} and the interaction energy expected from calibration sources. T_{eff} was calibrated to represent the kinetic energy of an electron created at the center of the detector *via* a small correction to E_{RSP} [62]:

$$T_{eff} = E_{\text{RSP}} \left(0.9850 + \frac{0.09693}{E_{\text{RSP}}} \right) - 0.511 \tag{2.4}$$

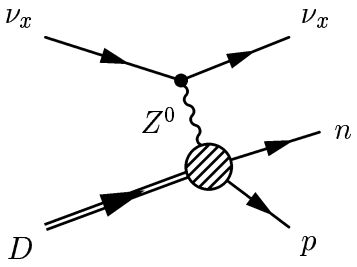
Throughout this thesis, an analysis kinetic energy threshold of $T_{eff} > 5.5$ MeV has been used.

Another energy variable referred to in this thesis it E_{gen}, the generated total energy of a simulated electron. Finally, E_ν defines neutrino energy.

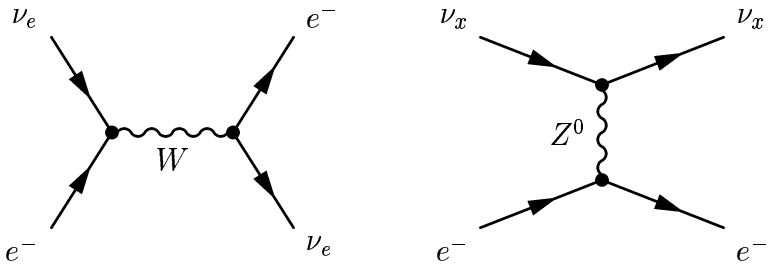
**This window relates to the time for a photon to travel directly to the PMT.



(a) Charged-current



(b) Neutral-current



(c) (d)

Elastic scattering

Figure 2.3: Diagrams of the neutrino interactions in SNO.

2.4.3 The Charged-Current Interaction

Charged-Current (CC) neutrino reactions in SNO result from W boson exchange between electron neutrinos and deuterons:

$$\nu_e + \mathrm{d} \rightarrow \mathrm{p} + \mathrm{p} + e^- \tag{2.5}$$

This interaction is shown in figure 2.3 (a) and is only sensitive to electron-type neutrinos. Due to the mass difference between the protons and electrons, the protons produced remain essentially stationary, and the electron carries off the initial neutrino energy minus a "Q value" of 1.44 MeV. The CC reaction therefore, also provides a measurement of the incoming neutrino energy spectrum. The reaction has a weak directional dependence with the distribution of the angle between the incoming neutrino and outgoing electron, $\theta_\odot$, given by [63]:

$$f(\cos\theta_\odot) = 1 - \frac{1}{3}\cos\theta_\odot \tag{2.6}$$

2.4.4 The Neutral-Current Interaction

The Neutral-Current (NC) reaction (shown in figure 2.3 (b)) is the neutrino disintegration of the deuteron:

$$\nu_x + \mathrm{d} \rightarrow \mathrm{p} + \mathrm{n} + \nu_x \tag{2.7}$$

This interaction has equal sensitivity to all active neutrino types and, therefore, can be used to measure the total neutrino flux from the Sun down to the reaction threshold of 2.22 MeV. The rate of this reaction will be unaffected by oscillations, unless oscillation to a sterile neutrino component occurs. The reaction can thus be used in conjunction with the CC flux measurement to provide a clear signature of neutrino oscillations.

The NC interaction is detected through the neutron produced. The SNO experiment was designed to run in three different phases, each with a different method for neutron detection. No information is obtained on the incoming neutrino energy in any of these phases as the signal arises solely from the energy released following neutron capture and bears no relation to the kinetic energy of the incoming neutrino.

In the first phase of the SNO experiment, which ran from November 1999 to June 2001, the detecting medium was pure D_2O. In this phase, neutrons were captured on deuteron nuclei, producing a 6.25 MeV γ. These subsequently Compton scattered electrons to energies sufficient to produce Čerenkov radiation.

In Phase II of the SNO experiment (the "salt phase"), 2 tonnes of NaCl were added to the D_2O volume. The salt provided Cl nuclei on which neutrons could capture with

a significantly larger cross-section. The capture efficiency was approximately 83% as compared with ≈30% corresponding to Phase I.

$$\mathrm{n} + {}^{35}\mathrm{Cl} \rightarrow {}^{36}\mathrm{Cl} + \Sigma\gamma \tag{2.8}$$

Neutron capture on deuteron releases a single γ, whereas capture on Cl typically releases a number of γs with a total energy of 8.6 MeV, each producing Compton scattered electrons observable through Čerenkov light. The increased energy release shifts the NC signal further away from low energy backgrounds, whilst the multiple γ production results in a more isotropic Čerenkov distribution which may be used to statistically distinguish neutron events from those which result in Čerenkov light produced by a single electron.

Following the removal of the salt via reverse osmosis, the third phase of the SNO experiment which will employ a method of neutron detection free from Čerenkov light, is currently being commissioned. Neutrons will be captured by ^{3}He proportional counters deployed throughout the D_2O volume. These Neutral Current Detectors (NCDs) will provide an independent measurement of the NC flux on an event-by-event basis. The neutron capture mechanism of these ^{3}He proportional counters is as follows:

$$\mathrm{n} + {}^{3}\mathrm{He} \rightarrow \mathrm{p} + \mathrm{t} + \gamma \tag{2.9}$$

Data from these three phases will supply independent checks of the total flux of active neutrinos produced in the Sun. The third phase will be subject to different systematic uncertainties to Phase I and II, which rely on statistical techniques to separate the CC and NC events. Combination of the data from these separate phases should substantially reduce uncertainties in the measured NC flux.

2.4.5 The Elastic Scattering Interaction

The elastic scattering (ES) reaction, given by equation 2.10, involves the scattering of neutrinos off electrons in both the D_2O and surrounding H_2O regions of the SNO detector.

$$\nu_x + \mathrm{e}^- \rightarrow \nu_x + \mathrm{e}^- \tag{2.10}$$

The reaction is sensitive to all three neutrino flavours, but has enhanced sensitivity to ν_e for which both W (figure 2.3 (c)) and Z (figure 2.3 (d)) boson exchange is allowed. For ν_μ and ν_τ neutrinos the reaction can only proceed through Z exchange and is approximately 6.5 times less sensitive than for ν_e. Due to conservation of momentum in the ES reaction, the distribution of electrons is strongly peaked in the direction away from the Sun.

2.4.6 Background Interactions

Radioactive backgrounds can mimic neutrino signals in SNO in a number of ways. These are briefly described in this section. Whilst every effort is made to reduce the levels of radioactivity in the SNO detector, the remaining backgrounds must be thoroughly characterised and either excluded from the data sample through analysis cuts, or taken into account in the signal extraction process. There are also backgrounds due to non-radioactive sources which are included here. The following discussion assumes the energy threshold ($T_{eff} > 5.5\,\mathrm{MeV}$) and fiducial volume (defined within a radius of 550 cm) that are used throughout this thesis.

Internal Photodisintegration Backgrounds

Neutrons can be produced from photodisintegration of deuterium nuclei by gammas of energy greater than 2.2 MeV, produced by products in the ^{238}U and ^{232}Th decay chains. These decay schemes are shown in appendix A. The numbers of ^{214}Bi and ^{208}Tl events were obtained from a maximum likelihood fit to isotropy distributions [49].

Neutrons can capture on ^{23}Na to produce an excited state of ^{24}Na, which has a half life of 14.96 hours and decays to produce βs and γs of sufficient energy to photo-disintegrate deuterium. Sodium, activated by calibration source neutrons, can decay in subsequent neutrino runs after the source has been removed from the detector. Neutrons entering into the D_2O from outside the detector can also produce ^{24}Na, especially in the neck region, and the D_2O recirculation systems, which have less shielding. ^{24}Na created in these regions and circulated into the fiducial volume must also be taken into account. Neutrons produced by muon spallation are removed from the data using a timing cut following muon signals (see appendix C), but their contribution to sodium activation is also included in the number of internal photo-disintegration backgrounds. ^{24}Na events have a similar distribution to ^{208}Tl events and are constrained in the fit described above.

A small number of neutrons may be produced by reactions of CNO cycle neutrinos, reactor neutrinos, terrestrial neutrinos and atmospheric neutrinos in the D_2O. Also (α,n) capture reactions on deuterium and ^{17}O or ^{18}O produce a small contribution to the neutron background.

Above the 5.5 MeV energy threshold the photodisintegration events are indistinguishable from neutral-current events and, therefore, must be subtracted from the fitted number of NC events.

Internal Čerenkov Backgrounds

β and γ radiation emitted from the ^{238}U and ^{232}Th decay chains and from ^{24}Na can also produce Čerenkov radiation. The SNO detector has limited energy resolution and upward fluctuations in the reconstructed energy can place these low energy events above the analysis threshold. This effect is small though, and the contribution of internal Čerenkov events is treated as a systematic uncertainty on the fit results. If the energy threshold is reduced in future spectral analyses, a more rigorous treatment of these background contributions will be required.

External β-γ Backgrounds

Radioactivity in the H_2O, AV, and PMT array will also produce βs and γs, which can be mis-reconstructed into the neutrino data set. The estimated contribution of these events is obtained from a fit between 1.1 and 2.5 cubed AV radii [64]. The contribution of these events to the neutrino data set is very small due to the selected energy threshold, fiducial volume and isotropy cuts.

External Neutrons

Photodisintegration neutrons can also be created in regions outside the D_2O. (α,n) reactions on ^{2}H, ^{17}O, and ^{13}C, which are abundant elements in the acrylic vessel, produce neutrons that can enter into the D_2O. Events due to these neutrons show similar trends to NC events in all observables except radius. The characteristic radial distribution for these events is determined from Monte Carlo simulation of neutrons in the AV.

High Energy γ Backgrounds

The highest energy of γ produced by the U and Th decay chains is 2.6 MeV, but γs of higher energy can be created in regions outside the D_2O. One mechanism is thermal neutron capture on high density materials, such as steel in the PSUP and cavity rock. The free neutrons originate from (α,n) reactions initiated by α emissions from the U and Th decay chains. γs can also be produced through reactions of the type:

$$\alpha + {}^{\mathrm{A}}_{\mathrm{Z}}\mathrm{X} \rightarrow {}^{(\mathrm{A}+1)}_{(\mathrm{Z}+3)}\mathrm{Y} + \mathrm{p} + \gamma, \tag{2.11}$$

where X are light nuclei such as Al, which is found in the PMT concentrators, and Y are the daughter products. These external γs can be produced with energies up to 10 MeV and therefore are not excluded by the energy threshold applied. However, since these events originate in the vicinity of the PSUP and cavity, they have a similar radial distribution to

external β-γ background events originating from the PSUP and, therefore, are included in the external β-γ background limit.

Other sources of γ background events include atmospheric neutrino interactions and the spontaneous fission of ^{238}U. The γ spectrum produced by atmospheric neutrino reactions, was estimated to contribute 7.3±2.2 γ events to the salt data set [65]. The spontaneous fission half life of ^{238}U is 8.2×10^{15} years, resulting in an estimated 12.72 ± 1.7 decays per year in SNO, of which only about 1 event enters into the salt data set [65]. As the number of events due to these background is small, they are simply included with a separate distribution of fixed amplitude in the signal extraction process. Such events are assumed to be evenly distributed throughout the D_2O volume and their energy distribution is approximated by simulated ^{16}N calibration events, *i.e.* γs of 6.13 MeV.

hep Events

Although SNO is predominantly sensitive to solar neutrinos produced by the ^{8}B reaction, it can also detect neutrinos from the rarer *hep* reactions.

$$^3\mathrm{He} + \mathrm{p} \rightarrow \ ^4\mathrm{He} + \mathrm{e}^+ + \nu_\mathrm{e} \tag{2.12}$$

The branching ratio for this reaction is $\ll 1\%$ (see figure 1.1). These neutrinos have a different energy spectrum from those produced by the ^{8}B reaction and could be considered as a background to the ^{8}B flux and spectrum measurements. A separate analysis is being developed by the SNO collaboration to measure the flux of *hep* neutrinos but, for the analyses presented in this thesis, the results presented contain combined contributions from neutrinos produced by both *hep* and ^{8}B reactions.

Figure 2.4 shows the predicted spectrum of CC events due to *hep* neutrinos. This spectrum was created from events simulated using the *hep* neutrino spectrum specified in [29], and has been normalised to the BP2004 [21] predicted flux of $7.88 \times 10^3\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for the data set analysed in this thesis. This can be compared to the predicted undistorted spectrum for ^{8}B CC events in the upper panel of figure 1.5 (which has been normalised to unity).

2.4.7 Other Background Events

The average data rate of the SNO detector is around 18 Hz, 5 Hz of which is due to the PulseGT. The majority of the events recorded are due to "instrumental backgrounds" defined as any trigger not associated with particle interactions in the detector [66]. For example, electronic pickup that can cause correlated discriminator firings. A series of data selection criteria, or "cuts" were developed based on information such as the average

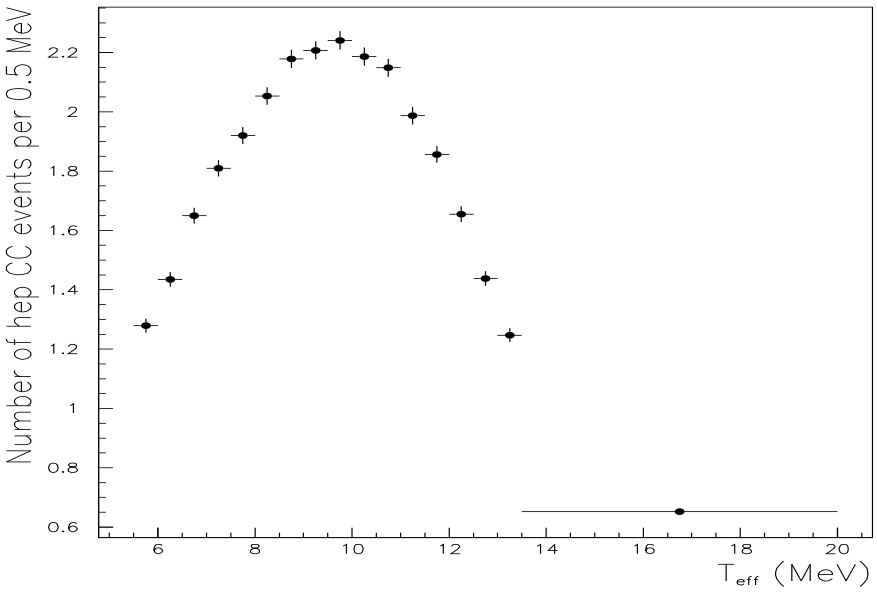


Figure 2.4: The predicted hep CC spectrum normalised to the number of events predicted by BP2004 to contribute to the data set analysed in this thesis.

event charge, clustering of hit PMTs, time correlations and raw time spectra. These cuts were applied in the first stage of data processing.

AV Events

One class of background events exhibits very isotropic hit patterns and appears to reconstruct at the AV. The exact cause of these events is unknown, but they may be associated with small regions of stress in the acrylic. These "AV events" were efficiently removed from the D_2O data set by an isotropy cut. This cut was relaxed for salt data analysis to avoid removing the more isotropic NC events from the data, but the contribution of AV events is still small enough to simply treat as a minor systematic uncertainty.

2.5 Summary

In this chapter the various subsystems which comprise the SNO detector have been discussed in some detail. A complete understanding of the detector response is required in order to obtain precise physics measurements from the SNO data. This is achieved through detailed simulations based on the findings of the rigorous detector calibration program. The next two chapters discuss details of the optical and energy calibration of SNO.

Chapter 3

Energy Scaling in SNO

A good understanding of the mapping between the observed energy deposited in the SNO detector and the true energy of the interacting particle is vital for the extraction of any meaningful spectrum. Uncertainties in this mapping will contribute to the systematic uncertainty of the spectral measurement. This chapter deals with an analytical calculation of the scaling between "true" energy and the number of PMTs hit ($Nhits$) in the SNO detector during the D_2O phase of the experiment.

3.1 Energy Calibration

Calibration source data are used to characterise the detector energy response, but only provide such calibrations at a number of fixed energy points. It is necessary to interpolate between these measurements in order to map from $Nhits$ to energy across the range of interest for neutrino interactions. Figure 3.1 shows the data and simulated $Nhits$ distributions for three different calibration sources: ^{16}N, ^{8}Li and Pt, which are described in table 2.2. The ^{16}N source of 6.13 MeV γs is the primary energy calibration source, whilst Pt provides another mono-energetic calibration point, producing 19.8 MeV γs. The ^{8}Li source produces βs with an energy spectrum extending up to about 14 MeV, similar to that expected for the CC signal (black stars), but the continuous spectrum does not constrain the energy scale as well as the mono-energetic calibration data.

3.2 Analytical Calculation of Energy Scaling

In order to calculate the energy deposited in SNO that will be observed, it is necessary to account for the following factors:

- Electron energy losses (including Bremsstrahlung emission and delta-ray production).

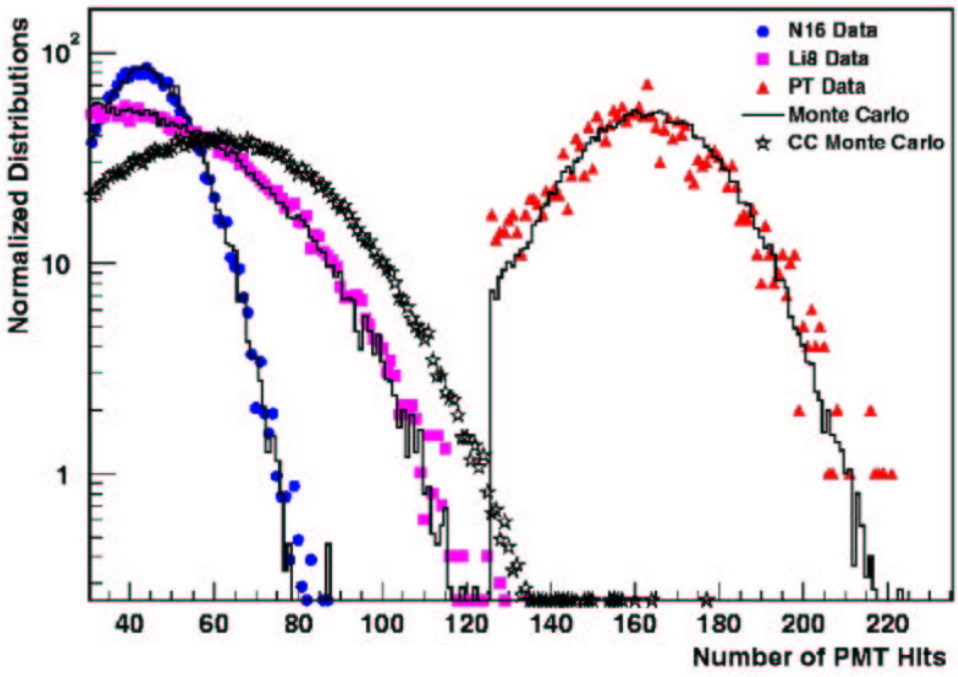


Figure 3.1: Calibration data and simulated events for ^{16}N, ^{8}Li and Pt data obtained during the D_2O phase. The expected CC distribution is also shown (black stars) to indicate the energy range of interest.

- Čerenkov light production
- **Multiple hits** - PMTs that are hit by more than one photon in the same event.
- **Noise hits** - Additional random hits due to PMT dark currents.
- **Cross-talk** - Electronic pick-up effects between adjacent electronic channels.
- **Energy drift** - Effects of time variation in detector response.
- **Detector Configuration** - The number and position of non-operational PMTs which changes over time.

Analytical calculation of these factors should give the correct relationship between energy and *Nhits*. A single data point from the ^{16}N source is then required to fix the parameterisation to account for the actual light detection efficiency of the detector.

A SNOMAN simulation of mono-energetic electrons indicated a linear scaling of *Nhits* with energy in the region of 5–10 MeV.* A similar result was obtained from a calculation of the expected light yield of Čerenkov events as a function of electron energy on an arbitrary scale using a numerical integration:

$$\text{Yield} = K \int \left[1 - \left(\frac{1}{\beta n} \right)^2 \right] dx \tag{3.1}$$

where K is an undetermined constant which involves the detector response, β, which is a function of x, is the electron velocity in units of c, and n is the refractive index of D_2O. Electron track length dx can be expressed in terms of kinetic energy T, with a parameterisation for the tabulated stopping powers of electrons in water, $\frac{dT}{dx}$ obtained from [67].†

$$dx = \frac{dx}{dT} dT \tag{3.2}$$

The parameterisation used is given in equation 3.3 and the fit to data is included in figure 3.2. The parameterisation is accurate to 0.5% in the region 0.25–20 MeV.

$$\frac{dT}{dx} = -\frac{1}{\beta^2} \left(1.649 + 0.18 \ln T + 0.0032 (\ln T)^4 \right) \tag{3.3}$$

The calculation was improved with a correction for *δ-rays* (high energy electrons which are periodically produced as the initial electron collides with matter). They cause secondary ionisation and produce their own Čerenkov light. An approximation was made

*The main region of interest for the observation of ^{8}B solar neutrinos in SNO.

†The tabulated numbers are for H_2O, not D_2O, but the difference in stopping power due to the higher density of D_2O is negligible.

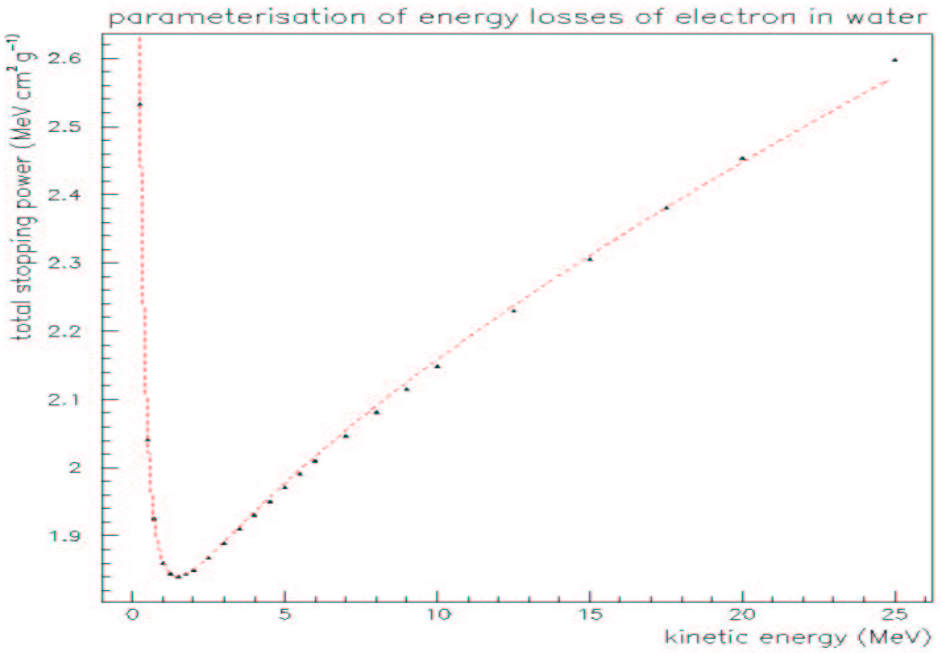


Figure 3.2: The parameterisation of equation (3.3) for the stopping power of electrons in water fitted to data taken from Ref. [67]

for the δ-ray path length, allowing for losses of $2\,\mathrm{MeV\,g^{-1}\,cm^{-2}}$ with a 0.5 MeV energy offset. The offset is due to the fact that the electrons slow down from speeds of $\beta \cong 1$ to the threshold for Čerenkov radiation.

$$l(T') = \frac{T' - 0.5}{2} \tag{3.4}$$

where T' is the δ-ray kinetic energy. In principle, one could use equation 3.3 to parameterise the further subsequent energy loss at this stage, but this significantly increases the complexity of the computation with relatively little gain in accuracy.

The probability of δ-ray production was approximated using the Rutherford cross-section (equation 3.5). Values for Z and A, the charge and mass numbers of the material, Avagadro's number, N_A, the classical electron radius, r_e, and mass, m_e have been inserted.

$$P(\delta) = \frac{2\pi N_\mathrm{A} Z r_e^2}{A} \frac{m_e}{\beta^2} \frac{dT'}{T'^2} = \frac{0.075}{\beta^2} \frac{dT'}{T'^2} \tag{3.5}$$

The resultant yield from δ-rays,

$$\mathrm{Yield}(\delta) = K \left(1 - \frac{1}{n^2}\right) \int_{0.5}^{T/2} \left(\frac{T' - 0.5}{2\,\mathrm{MeV}}\right) \left(\frac{0.075}{(T')^2 \beta^2}\right) dT', \tag{3.6}$$

must be added to the initial light yield before the integration of equation 3.1. The maximum value of T' is half that of the initial electron, which provides the limits for this integration.

A correction was also made for Bremsstrahlung emission, which decreases the light yield at higher initial electron energies. To calculate this effect, Bremsstrahlung photons of energy E_{B} were approximated as contributing the Čerenkov radiation of an electron with initial speed, $\beta = 1$, and energy $E_{\mathrm{B}} - 1.5\,\mathrm{MeV}$. 1 MeV is subtracted to account for the difference in the amount of Čerenkov light obtained from an electron and a γ, and 0.5 MeV is subtracted because the electron slows down (as for δ-rays).

With the Bremsstrahlung and δ-ray corrections, both of which have roughly 5% effect at 20 MeV, the result is surprisingly linear, taking the form of equation 3.7.

$$\text{Yield} \propto (T - 0.45) \tag{3.7}$$

This is also in good agreement with past calculations made using the EGS4 electron and gamma transport package [68]. The result is plotted in figure 3.3.

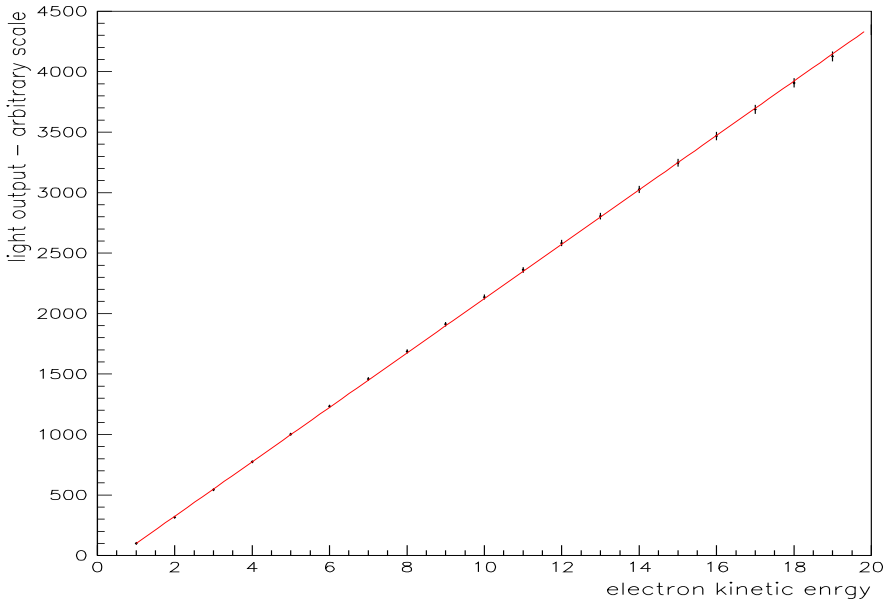


Figure 3.3: A plot showing the linear scaling of light yield with electron kinetic energy when both δ-rays and Bremsstrahlung emission are taken into account. An uncertainty of 1% is assumed for each point, due to the assumptions made within the calculation. The fit takes the form of equation 3.7.

3.3 Fixing the Scale

The scale was then fixed using data from the ^{16}N calibration source deployed at a variety of positions inside the detector. A volume-weighted average of the *Nhits* value for this source was obtained over a number of runs and found to be 47 hits, accurate to 1% [69]. A SNOMAN simulation of 6.1 MeV gamma rays was found to produce the same mean *Nhits* as an electron of kinetic energy 5.03 ± 0.02 MeV. The error quoted here is the statistical uncertainty derived from the spread of the distribution. This result was verified by an analytical calculation of the light radiated in multiple Compton Scattering [70]. This calculation showed an electron of energy 5.1 MeV produced the equivalent light of the ^{16}N γ, which is in agreement at the 1% level.

The mean *Nhits* value from ^{16}N accounts for the detector response efficiency in the calculated scale. However, it is also necessary to consider other effects which may not have a linear contribution over the energy range in question.

3.4 Multiple Hits

A PMT can only fire once for any event, so any secondary photons which also create photo-electrons in that PMT will not contribute to *Nhits*.[‡] The relevance of this effect will be greater for events with higher light yields and hence, a larger overall value of *Nhits*. The effect is also dependent on the position and direction of an event.

The most probable number of multiple hits for an event of given *Nhits* was determined from the average number of photons impinging on each PMT, μ, which depends on the detector geometry and angular distribution of Čerenkov light. The distribution of Čerenkov light from ^{16}N, ^{8}Li and neutrino data was obtained by considering the angle of each struck PMT with respect to the reconstructed event direction. These data are included in figure 3.4, along with a theoretical calculation of the Čerenkov light distribution from electrons with a kinetic energy of 5 MeV [71]. The theoretical calculation averaged the angle of emitted photons relative to the true electron direction over many electrons. The increased peak in the data with respect to the theoretical calculation is thought to be partly due to the way the angle is effectively quantised by the finite size of the PMTs. A slight reconstruction bias may also contribute to this effect as the reconstruction algorithm[§] used determines track direction by maximising the symmetry of the Čerenkov ring.

[‡] As photon detection efficiency is significantly less than 100%, multiple hits are defined as photons that would have fired the PMT if it had not already fired.

[§] The PATH fitter [72] which uses both timing and angular PMT information was used to reconstruct all the data discussed in this chapter.

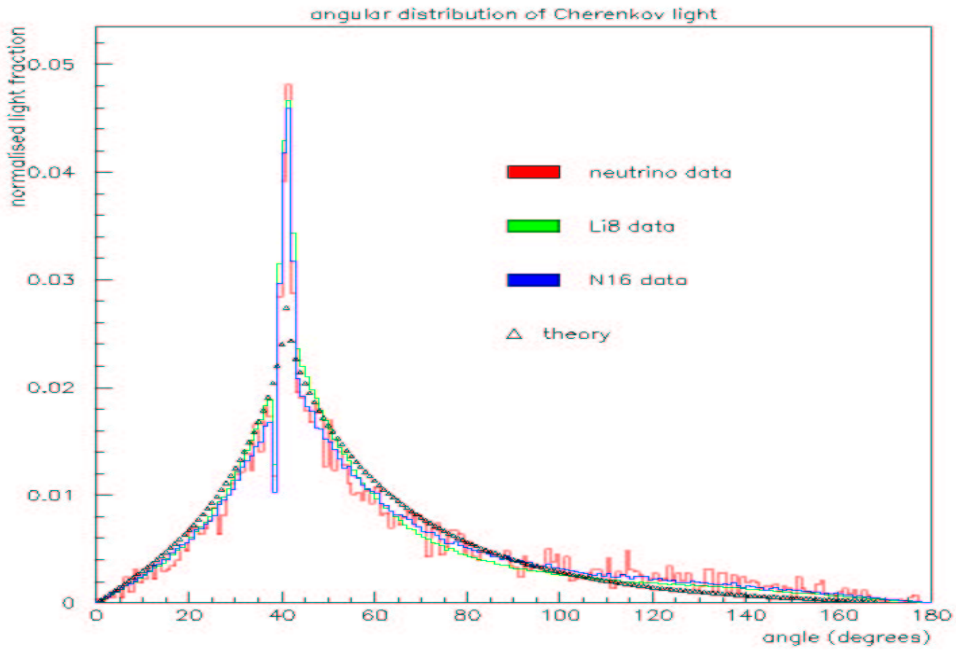


Figure 3.4: Angular distribution of Čerenkov light in the SNO detector as measured with various calibration sources. A theoretical distribution calculated for 5 MeV electrons is also shown. The peak is at $\approx 42°$, the half angle of the Čerenkov cone in D_2O.

To calculate μ, the angle between each PMT and the track direction was determined using geometry information recorded in SNOMAN. The relative fraction of Čerenkov light observed at this angle, $f_{\rm den}$, was extracted from an array containing the data in figure 3.4 and weighted by the solid angle subtended by the PMT at the event vertex.[¶] The value was then approximately normalised by the number of hits observed:

$$\mu \cong Nhits \times f_{\rm den} \times \frac{\Omega_{\rm PMT}}{4\pi} \tag{3.8}$$

In this scenario, 100% coverage of PMTs[‖] was assumed since the detector response was taken into account by normalising to the observed number of hits. This was significantly easier than normalising to the number of photons produced, which would require a knowledge of the photon detection efficiency. For events at non-zero radii, the solid angle was calculated separately for each PMT.

For a given event position and direction, Poissonian statistics were used to calculate the probability of a PMT firing multiple times for a given μ. The probability that a PMT will fire was obtained from a series summation for any number of PMT hits:

$$\sum_{n=1}^{\infty} P(n/\mu) = 1 - P(0/\mu) = 1 - e^{-\mu} \tag{3.9}$$

where $P(n/\mu)$ is the probability that n photo-electrons are produced in a PMT that is hit by μ photons.

The measured $Nhits$ for an event was equated to the sum of equation 3.9 over all PMTs in the detector, $N_{\rm PMT}$, as a single PMT hit is recorded for each PMT in which ≥ 1 photo-electron has been produced.

$$Nhits = \sum_{i=1}^{N_{\rm PMT}} (1 - e^{-\mu(i)}) \tag{3.10}$$

The summation of μ over all PMTs in the detector approximates the total light deposited in the detector. The total number of multiple hits, $N_{\rm MH}$, for any event is obtained from equation 3.11.

$$N_{\rm MH} = \sum_{i=1}^{N_{\rm PMT}} \left(\mu(i) - \left(1 - e^{-\mu(i)}\right)\right) \tag{3.11}$$

[¶] Rather than take the physical radius of a concentrator to obtain the solid angle, the total PSUP surface area was divided by the number of PMTs to obtain the area covered by each PMT assuming no gaps between them. The radius this area would have if it were circular was then used for the calculation of solid angles. This effectively gives 100% coverage of PMTs.

[‖] The actual PMT coverage is ≈58%.

The measured $Nhits$ value for each event was corrected by the factor C_{MH} (equation 3.12) to obtain the effective amount of light in the detector.

$$C_{\mathrm{MH}} = 1 + \frac{N_{\mathrm{MH}}}{Nhits} \tag{3.12}$$

An overall correction factor applicable to isotropic neutrino data, C_{MH} was obtained by averaging over events distributed out to an assumed fiducial radius of 5.5 m and over all track directions. The resultant percentage correction for a range of measured $Nhits$ values is included in figure 3.5 and is well fitted by the polynomial:

$$C^{\mathrm{iso}}_{\mathrm{MH}}(\%) = 0.011 + 0.0206 Nhits + 0.000022 Nhits^2 \tag{3.13}$$

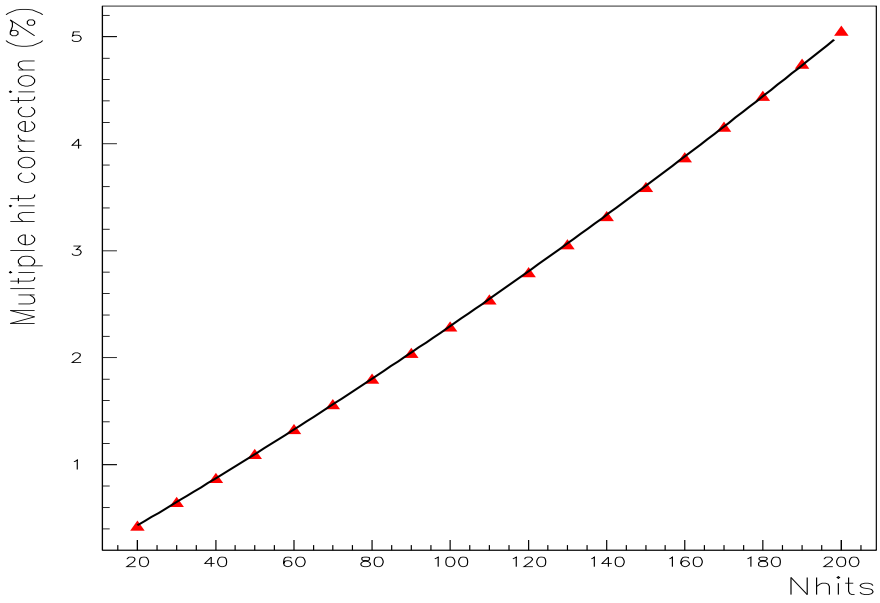


Figure 3.5: Percentage correction for isotropic data to account for multiple hits, fitted with a second order polynomial.

This correction was used to translate the predicted effective light level from the analytical calculation to $Nhits$. A similar technique was used to determine the apparent reduction in observed energy due to multiple hits from calibration data by considering localised events at a given point in the detector.

3.5 Other Effects

The effect of multiple hits is, in part, counteracted by the phenomenon of cross-talk. A high charge output from a PMT (which is more likely if the PMT is hit multiple times) increases the chance of inducing hits in neighbouring electronic channels that result in an additional "hit". An empirical parameterisation of cross-talk as given in equation 3.14 was used to include this effect, where N_{Xt} is the number of hits gained through cross-talk [73]:

$$N_{\mathrm{Xt}} = 0.94 + 0.013 \times (Nhits - N_{\mathrm{Xt}}) \tag{3.14}$$

The measured $Nhits$ for an event was multiplied by the factor C_{Xt} to remove the contribution from cross-talk hits.

$$C_{\mathrm{Xt}} = 1 - \frac{N_{\mathrm{Xt}}}{Nhits} \tag{3.15}$$

Random noise hits were taken into account by subtracting 2 hits per event. This is the average number of random noise hits occurring in the event trigger window, as established from data obtained from the Pulse GT.

During the D_2O phase, a change in the detector response was observed over time. This "energy drift" was quantified by a reduction in $Nhits$ of 2% per year[74].

3.6 Tests of the Scaling Calculation

This theoretical model was tested by extrapolating from the ^{16}N energy to the energy of the Pt source. A SNOMAN simulation was used to infer that the light produced by the Pt γ is on average, equivalent to an electron with kinetic energy 18.6 ± 0.5 MeV. This was also verified by numerical calculation which gave the equivalent mean electron energy to be 18.5 ± 0.2 MeV [70]. The average energy loss in creating electrons is larger for the Pt γ than for ^{16}N γs due to the higher probability of electron pair production at these energies, which results in the loss of visible light to the electron rest mass energy.

Finally, the angular distribution of events from calibration sources was considered. The ^{16}N source emits events nearly isotropically but the Pt source emits γs equatorially into a $\sin^2\theta$ distribution [75]. This can be seen in the angular distribution of hit PMTs in the Pt data. Separate multiple hit corrections were calculated for each source run, taking into account the source position and the angular distribution.

Figure 3.6 shows the results of extrapolation from the average $Nhits$ obtained from data taken with both the ^{16}N and the Pt source run at the centre of the detector. The procedure for the prediction of Pt source $Nhits$ was as follows.

1. First, the measured $Nhits$ value obtained from ^{16}N data was reduced by a factor that accounted for both the energy drift between the time of ^{16}N and Pt data taking and the change in the number of operational PMTs.

2. This scaled ^{16}N *Nhits* value was then multiplied by C_{MH}, the multiple-hits correction and C_{Xt}, the cross-talk correction.

3. The linear parameterisation given in equation 3.7 was fixed to this "corrected" ^{16}N *Nhits* value and extrapolated to Pt energies.

4. Finally, the predicted value at the Pt energy was divided by C_{MH} and C_{Xt} in order to compare it to the measured Pt *Nhits* value.

In figure 3.6 the Pt point is the true data measurement but the ^{16}N value was scaled by 1.0488 to account for energy drift and the reduction in operational PMTs. The line gives the number of hits predicted for a given energy at the time of the Pt data taking.

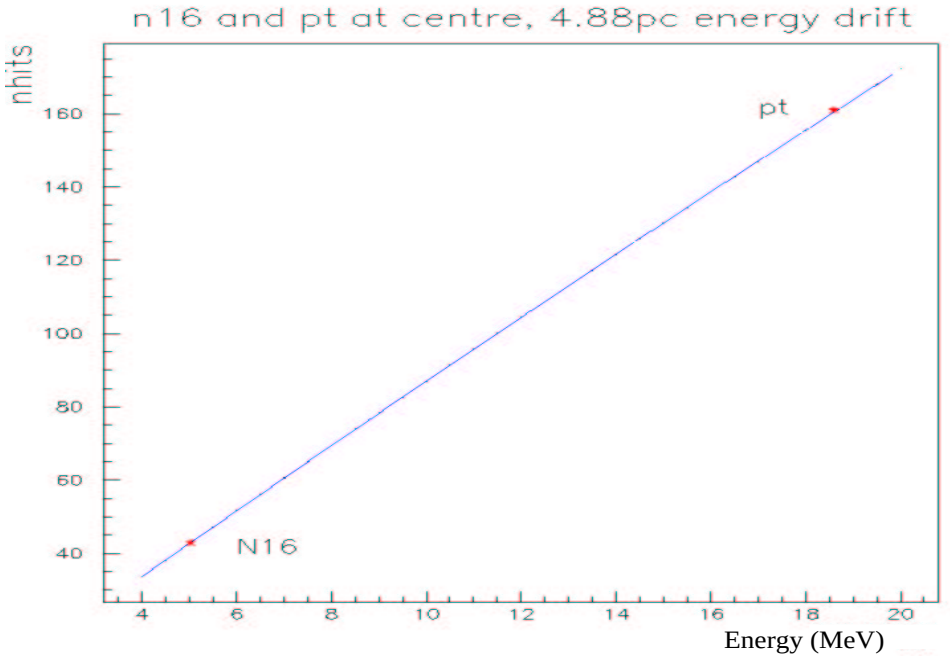


Figure 3.6: Theoretical energy scaling of *Nhits* for sources at the centre of the detector. The error bars on the 2 source data points are too small to be visible. The 4.88% energy drift includes a correction for the number of tubes online at the time of data taking and has been applied to the ^{16}N *Nhits* value.

Further tests were made on off-axis data, with the calibration sources situated at positions 2.5 m, and 5 m below the centre of the detector. The results of these tests are shown in figures 3.7 and 3.8 respectively.

The error on the average *Nhits* was $\pm$ 0.05 for ^{16}N and $\pm$ 0.5 for Pt data. Table 3.1 summarises the values for the average measured *Nhits* for Pt and those predicted by the extrapolation.

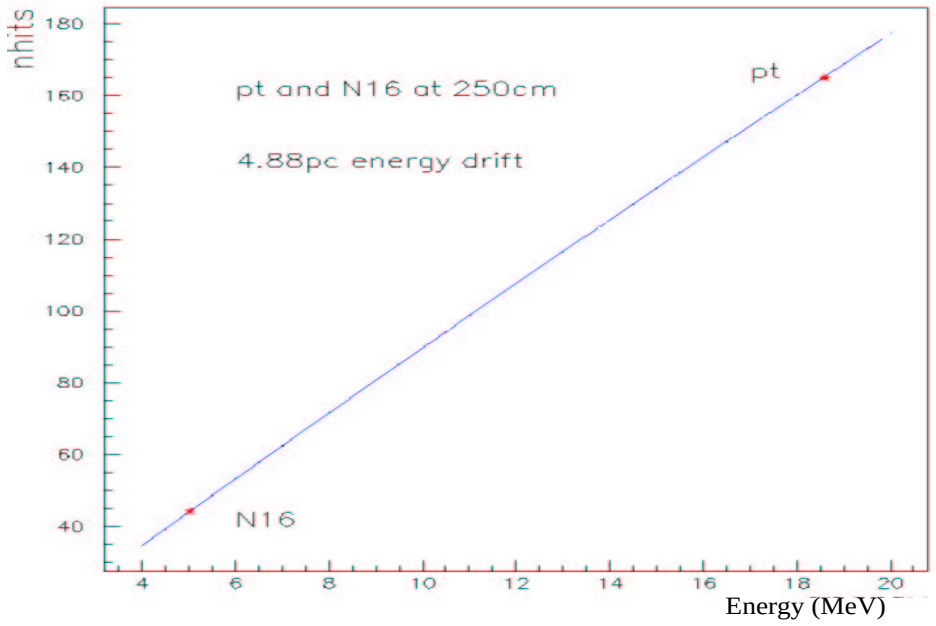


Figure 3.7: Theoretical energy scaling of *Nhits* with the Pt source at $z = -250\,\text{cm}$. The average *Nhits* for runs at $z = -300\,\text{cm}$ and $z = -200\,\text{cm}$ are taken for ^{16}N source. The 4.88% energy drift includes a correction for the number of working PMTs at the time of data taking and is applied to the ^{16}N *Nhits* value as in figure 3.6.

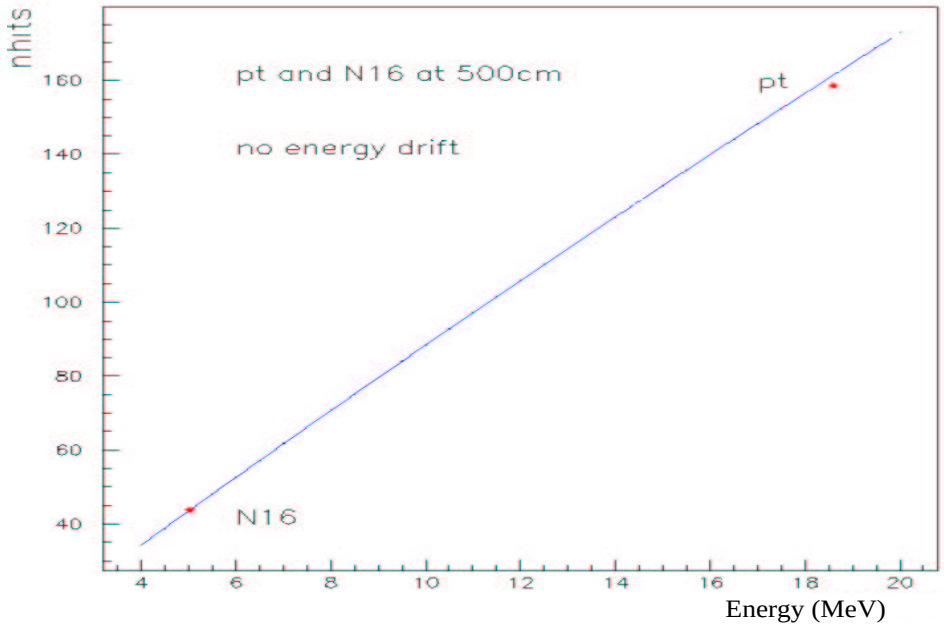


Figure 3.8: Theoretical energy scaling of *Nhits* for calibration sources deployed at $z = -500\,\text{cm}$. There is no energy drift in this plot as the ^{16}N run used was taken close to the Pt run in time.

Position	Pt average *Nhits*	Predicted *Nhits*	percentage discrepancy
centre	161.0 ± 0.5	159.7	0.81% (Nov 1999 ^{16}N)
-250cm	165.0 ± 0.5	164.5	0.30% (Nov 1999 ^{16}N)
-500cm	158.6 ± 0.5	165.4	4.27% (Nov 1999 ^{16}N)
-500cm	158.6 ± 0.5	160.6	1.26% (Jan 2001 ^{16}N)

Table 3.1: Summary of the agreement between calculated (predicted) *Nhits* at 18.6 MeV, and the average value measured for the Pt source.

It should be noted that although the number of non-operational PMTs was taken into account, the position of these PMTs was not considered. For runs at larger radius, this appears to have more impact. Two ^{16}N runs at $z = -500$ cm were available for comparison to Pt data, one in November 1999 and one from January 2001. Results of scaling from both are included in table 3.1. The difference between the Pt data and the November 1999 ^{16}N run at $z = -500$ cm is significantly greater at 4.27%, reflecting that changes in the detector configuration over time have more effect at larger radii.

3.7 Summary

As has been shown, it is possible to analytically extrapolate from a single measured value from the ^{16}N source, across a range of 13.7 MeV, to predict the *Nhits* value of the Pt source to an accuracy of better than 1%. This is good evidence that all appreciable detector effects, which may not scale in a linear fashion with energy, have been included. This independent check demonstrates a good understanding of the energy scaling in the SNO detector, and provides confidence in the simulations used to model events which also predict a linear relationship between generated energy and *Nhits* (see figure 3.9). The analyses described in the following chapters rely heavily on the ability of these simulations to accurately model the response to neutrino signals.

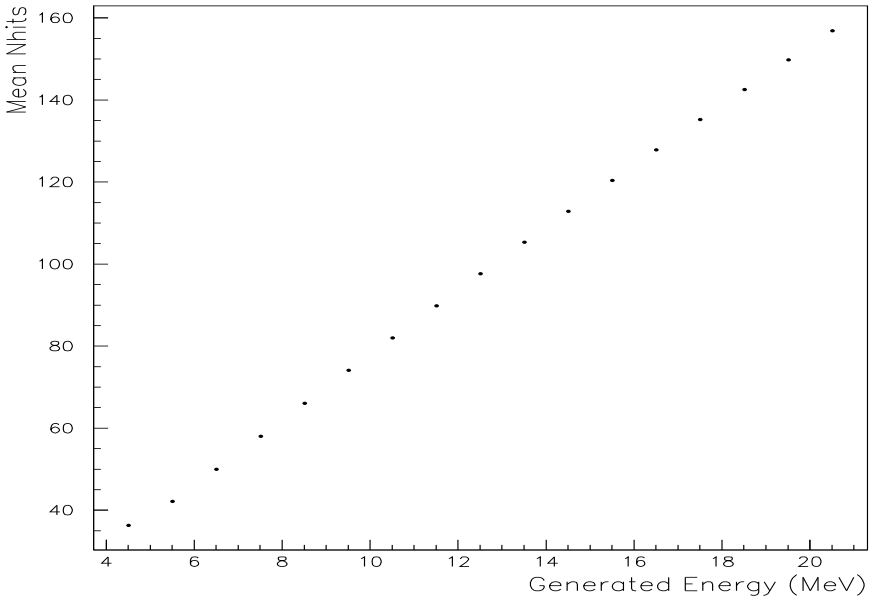


Figure 3.9: Measured *Nhits* for simulated electron events with a range of generated energies, modelled by the SNOMAN Monte Carlo package.

Chapter 4

Measurement of Rayleigh Scattering

To analyse SNO data a full understanding of the detector's optical response is required. One aspect of the optical calibration program for the salt phase was the measurement of the level of scattering in the heavy water. Scattering of photons in the detector can impact on the determination of many observable parameters. A poor understanding of the observable parameters used to separate the different event classes will lead to a bias in measured fluxes. In this chapter the first direct measurement of D_2O scattering will be presented.

4.1 The Effect of Scattering on SNO Measurements

One observable that can be affected by scattering is the measured event energy, T_{eff}, as photons that undergo scattering are delayed, and may cause PMT hits outside the time window used for analysis. Figure 4.1 shows the change in the T_{eff} distribution for CC events when the amount of Rayleigh scattering in the D_2O is changed by a factor of two. The predicted change in this distribution is comparable in magnitude to that caused by MSW distortions (see figure 1.5).

The timing of PMT hits in an event is also used to reconstruct the interaction position within the D_2O volume. Incorrect measurements of the amount of light delayed by Rayleigh scattering could bias the mapping of event position.

The angular distribution of observed light can also be affected by scattering. This is of particular importance in the salt phase as a parameterisation of event isotropy, β_{14} (discussed in more detail in appendix D) was one of the observables used to separate CC and NC events. Figure 4.2 shows the change in the β_{14} distribution for CC events when the level of Rayleigh scattering in the D_2O is changed by a factor of two. Events become more isotropic (lower β_{14}) as scattering is increased.

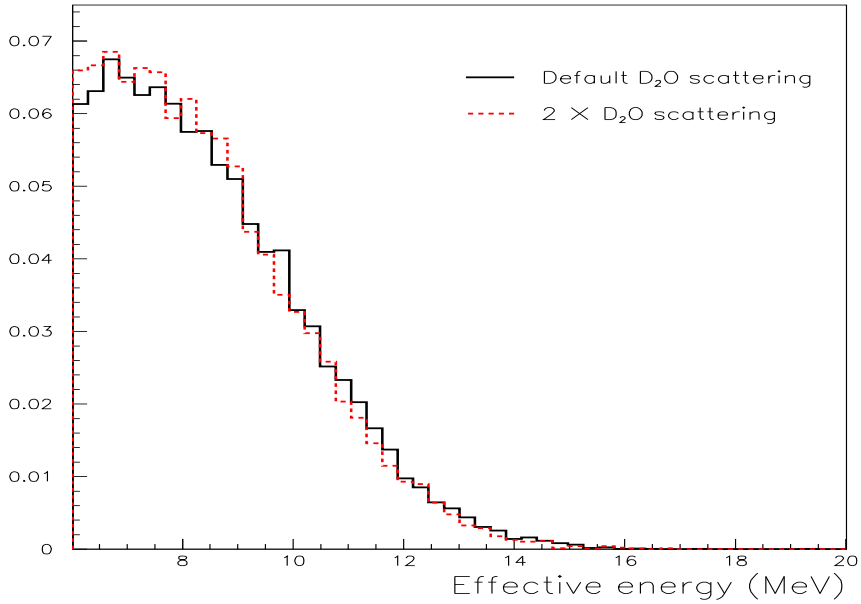


Figure 4.1: The effective energy distribution for simulated CC events with theoretical levels of Rayleigh scattering (black solid) and double Rayleigh scattering (red dashed) in the D_2O.

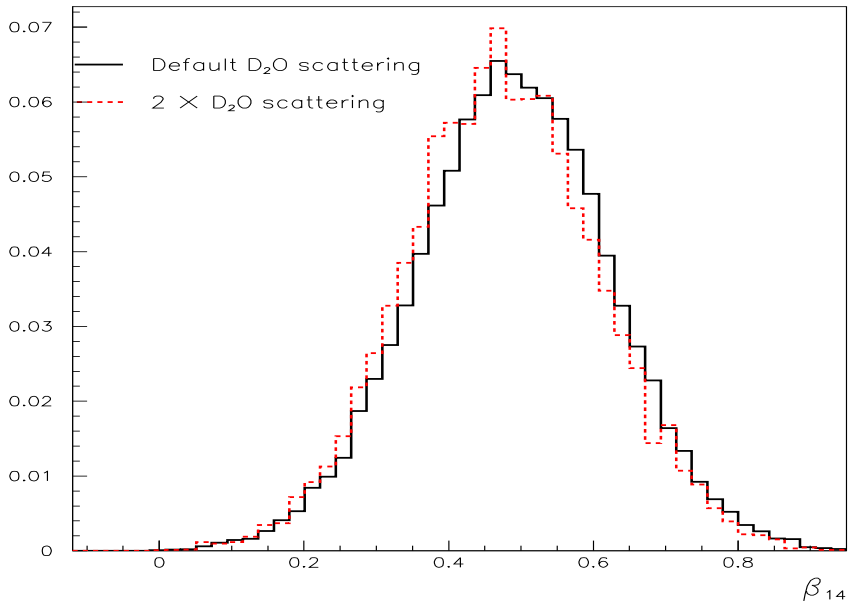


Figure 4.2: The isotropy distribution for simulated CC events with theoretical levels of Rayleigh scattering (black solid) and double Rayleigh scattering (red dashed) in the D_2O.

4.1.1 Possible Causes of Increased Scattering

A number of factors could increase the levels of scattering in the SNO detector. If not controlled, a variation in the temperature levels in the D_2O region could result in a change in the true Rayleigh scattering. Contaminants in the water can also cause Rayleigh-like scattering. One possible candidate is organic matter which could enter the detector from the water purification systems [76]. Alternatively, other materials could have been introduced to the detector along with NaCl, or perhaps in the process of deploying calibration sources.

4.2 Types of Scattering

There are three modes of scattering to consider in the SNO detector. The first is Rayleigh scattering, a phenomenon that occurs for particles of sizes $\leq \frac{\lambda}{15}$ [77], (where λ is the wavelength of incident light) such as atoms and molecules which have diameters of a few tens of nm. Incident photons induce vibrations in the electron cloud of such particles, setting up an electric dipole moment that immediately re-emits radiation at the same wavelength. In a gas the molecules are randomly distributed and there is no sustained pattern of coherence between the secondary wavelets produced so scattering can be observed at large angles. Rayleigh scattering, averaged over all possible polarisations, has an angular distribution of the form $(1 + \cos^2\theta)$, where θ is the angle between the initial and scattered photon. In liquids, the molecules are in a more ordered state so the amount of scattering per molecule is suppressed with respect to gases, although density fluctuations on a local scale serve to increase the scattering effect. As molecules have more electronic resonance states in the UV, the amplitude of vibration increases with frequency, and Rayleigh scattering scales with λ^{-4}.

The level of true Rayleigh scattering in the water of the SNO detector has been theoretically calculated [78, 79, 80] based on precise measurements of the temperature of the different regions. The water regions of the SNO detector are kept at 11 ± 2°C. The scattering lengths calculated for the D_2O region which were used as default input values for optical simulations are given in table 4.1.

Mie scattering, which occurs off spherical bodies, is generally independent of wavelength for particle diameters $\geq \lambda$. It could be caused by foreign impurities such as bacteria in the water of the SNO detector. Rayleigh scattering can be considered as the small-size limiting case of Mie Scattering.

The last type of scattering is specular scattering which could be caused by large particulate contaminants in the SNO detector.

Previous investigations indicate that the SNO detector is only sensitive to wide

Wavelength (nm)	Rayleigh scattering length (m)
337	75.69
365	107.36
421	198.50
500	410.71

Table 4.1: The theoretical Rayleigh scattering lengths for D_2O in SNOMAN.

angle scattering, including both Rayleigh scattering and specular scattering. This can be parameterised by a single effective scattering length [81]. Thus a calibration source and analysis were developed for the purpose of measuring the *effective Rayleigh scattering* in the D_2O region of the SNO detector.

4.3 Existing Optical Calibrations

The main tool for optical calibration of the SNO detector is the *laserball*: a pulsed laser source that provides monochromatic light at one of six different wavelengths (337, 365, 386, 421, 500 and 620nm). The laser light is fed down a fibre optic cable to the source. The light is diffused in a ball containing a mixture of silicone gel and small glass beads (with average diameter 50 μm) to try to produce an isotropic source of light.* The rate of laser pulses is variable, as is the intensity, which is adjusted using neutral density filters placed in the laser beam. The source can be deployed at a range of positions in the D_2O and H_2O volumes of the detector.

Optical parameters are extracted from calibration scans in which the laserball source is placed at many different positions in the detector. The photon path length and angle of incidence of photons on each PMT changes with source position and thus allows various optical parameters to be fit. The required parameters are the angular response of the PMTs (and concentrators), scattering and attenuation in the AV, reflections, and attenuation in both the D_2O and H_2O regions.

The fitted attenuation, $\alpha_{\rm net}$, is the combined effect of absorption and scattering in the water. However, to simulate events in the SNO detector, these phenomena must be treated separately, and both absorption and scattering lengths are required as inputs to the Monte Carlo. Equation 4.1 parameterises the relative contribution of absorption and scattering, which also depends on the amount of scattered light arriving in the prompt time window, k.

$$\alpha_{\rm net} = \alpha_{\rm abs} + k\alpha_{\rm scat} \tag{4.1}$$

*A slight anisotropy which remains in the source is modelled in the SNOMAN Monte Carlo.

The factor k is determined by geometric effects and should be independent of wavelength for true Rayleigh scattering. If all scattered light were removed from the time window used for the optical analysis then $k = 1$, but this is not the case due to the forward scattering.

4.4 Hardware Design

To directly measure Rayleigh scattering in SNO, a directional source was created. This confined prompt light to a limited region of the detector establishing an analysis region predominantly sensitive to scattered photons. The fraction of scattered light was determined from the amount of light initially deposited in the detector, and the number of photons scattered into this selected region.

Since the properties of the laserball (such as the pulse time structure) are generally well understood, a mask was constructed to produce a beamed source of light from the existing laserball rather than designing a completely new directional source. The mask was designed so that the beam pointed vertically downward due to the symmetry of the SNO detector about the central axis and the nature of the source deployment system.

4.4.1 Beam Width

The choice of beam width was influenced by two factors: a narrow beam was preferable to allow tight cuts on the prompt beam light and reflections, but the beam had to be wide enough to illuminate enough PMTs to measure the amount of light deposited in the detector.

The frequency of hits on the PMTs directly illuminated by the source (PMTs in the *beamline*) was limited to avoid the phenomenon of multiple hits (described in 3.4). The fraction of PMTs in the beamline that are hit in any given event, $F_{\rm bh}$, was limited to $\lesssim 10\%$, which also prevented saturation of the small number of PMTs at the bottom of the SNO detector.

To probe the scattering phenomena thoroughly, data were required for several wavelengths over a number of path lengths.[†] To obtain enough data in two 8 hour shifts, each run was limited to 15–30 minutes in duration.

A suitable beam width was determined by a rough calculation of the minimum number of PMTs in the beamline, $N_{\rm min}$, using the formula:

$$N_{\rm scat} = N_{\rm min}\mathcal{R}tF_{\rm bh}\eta\frac{P}{\Lambda} \tag{4.2}$$

[†]The path length in question is the distance travelled through the D_2O by photons detected in the beamline.

Part	Material
Can	Stainless steel
Upper sliding plate	Stainless steel
Lower sliding plate	Black polypropylene
Shielding cone piece 1	White polypropylene
Shielding cone piece 2	White polypropylene
Main collimator piece	Black polypropylene
Collimator end piece	Black polypropylene
Collimator baffling	Black polypropylene
Collimator insert disk	Black polypropylene
2 collimating washers	Stainless steel

Table 4.2: Table of the mask components and materials.

in which $\Re$ is the pulse rate, taken as 45 Hz (the maximum rate for the laserball source), P is the path length, Λ is the Rayleigh-like scattering length, η is the fraction of scattered light in the analysis region (estimated to be 50%), t is the run duration and $N_{\rm scat}$ is the number of hits in the analysis region. $N_{\rm scat}$ is subject to statistical uncertainty, which limits the precision of measurements. A value of $N_{scat} \geq 2500$ has $\lesssim 2\%$ uncertainty, considered a suitable limit for this measurement. The calculation was carried out for an 8.5 m path length and a run duration of 30 minutes where $F_{\rm bh}$ was taken to be 10%. For wavelengths of 400 nm a value of $N_{\rm min} = 30$ was obtained.[‡] For this path length, it was determined that a beam subtending a half angle of 7.0° would directly illuminate 32 PMTs. This is therefore a suitable value for the half-angle of collimation for data collected over path lengths of $\geq$8.5 m.

4.4.2 Mechanical Issues

Figure 4.3 shows the general assembly of the laserball mask. The material used for each part of the mask is given in table 4.2. Figure 4.5 is a photograph of the fully assembled mask, whilst figure 4.4 shows the separate pieces of the mask.

The design consisted of a number of sections and was developed by the author to satisfy the following criteria.

- **Size**: Dimensions of the mask were tightly constrained by the diameter of the laserball and the 12.7 cm diameter of the guide-tube through which the sources

[‡] As Rayleigh scattering scales with λ^{-4}, a long wavelength with respect to the mean wavelength of Čerenkov light was used for the calculation in order to obtain a lower limit on $N_{\rm scat}$.

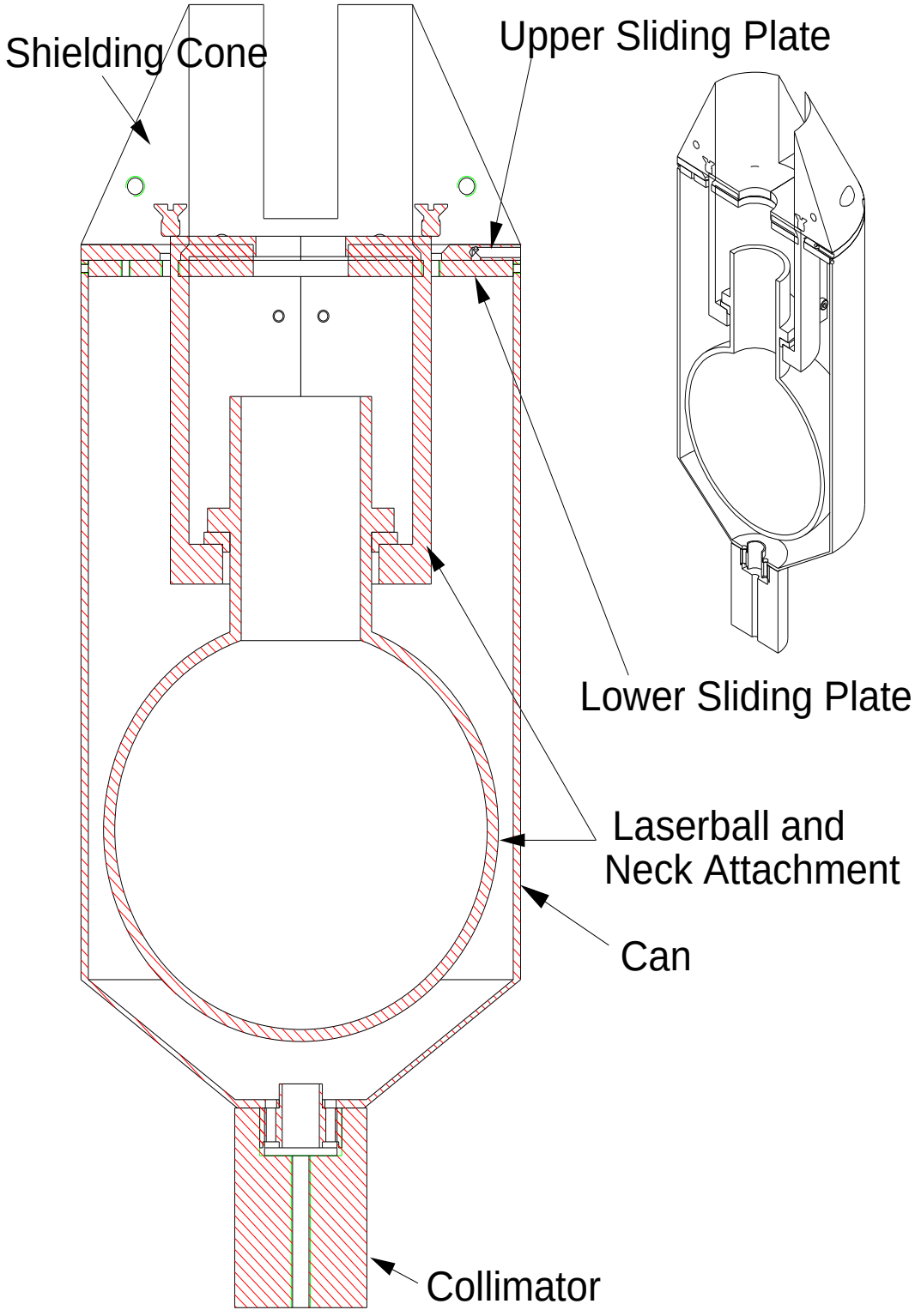


Figure 4.3: Diagram of the laserball mask, assembled around the laserball.

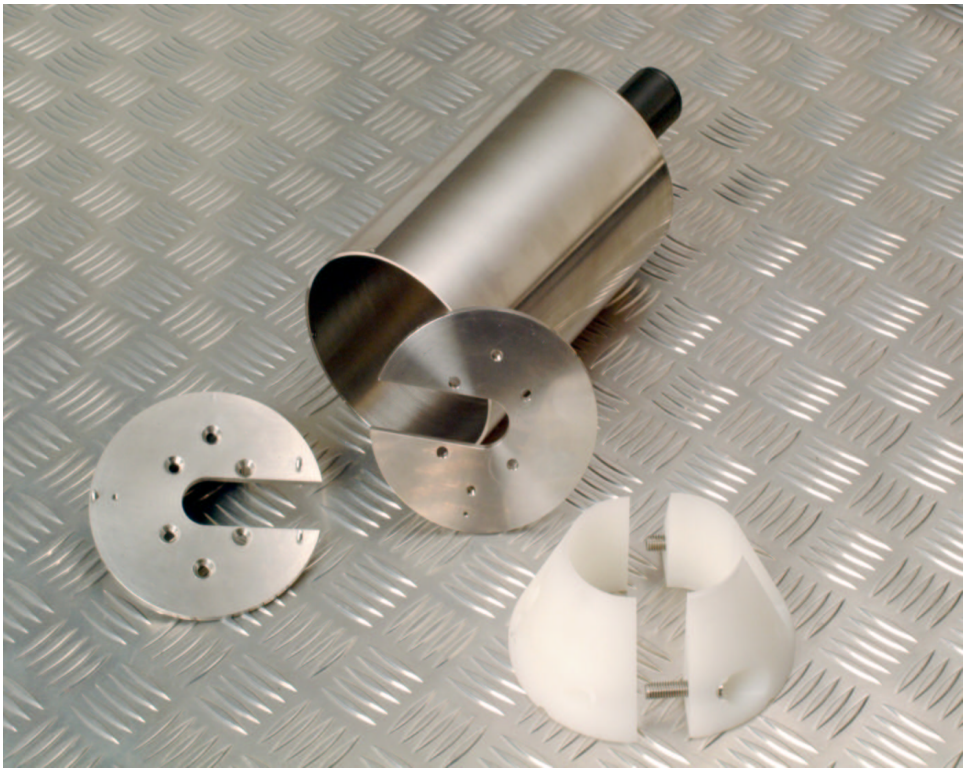


Figure 4.4: Photograph of the laserball mask parts taken at the University of Oxford. The main can piece, upper and lower sliding plates and neck shield are shown here.

Figure 4.5: Photograph of the assembled laserball mask taken at Queen's University.

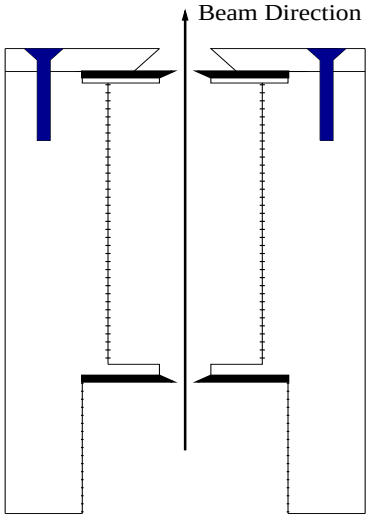


Figure 4.6: Design of the collimator for the laserball mask. As drawn, the bottom screws onto the base of the "can" and the collimated beam emerges from the top. The black polypropylene parts are shown in white, whilst the stainless steel collimating disks are shown in black. (The screws are shown in navy blue.)

enter the SNO detector.

- **Collimator design**: Figure 4.6 shows the design for the collimator, developed to give a sharp beam. Black polypropylene, roughened on the inside, and sharp edged metal disks prevent large angle reflections from the aperture.

- **Light seals**: All the seals were designed to be light-tight to the single photon level.

- **Deployment**: A white polypropylene cone was incorporated in the design to prevent sharp edges hampering deployment of the source. Drainage holes allow water to enter and leave the mask in a timely fashion, without compromising the light seals.

- **Attachment to the laserball**: The mask was designed to attach to the laserball when in position on the deployment mechanism.[§] Figure 4.4 shows sections of the mask prior to attachment around the laserball.

- **Structural stability**: There must be no danger of any piece of the mask falling off inside the detector or getting stuck on the laserball. In figure 4.4, the lower sliding plate is made of stainless steel. In the ultra-clean environment of the SNO laboratory, stainless steel screws binding into a drilled hole in stainless steel are easily

[§]This minimises the amount of neutrino data loss, due to activity above the detector.

Separation (cm)	Diameter illuminated (mm)	calculated opening angle(°)
13.5 ± 0.2	31 ± 1	11.5 ± 0.4
15.0 ± 0.2	36 ± 1	12.1 ± 0.4
85 ± 1	175 ± 10	11.4 ± 0.7

Table 4.3: Results of measurements of beam dispersion using photographic paper. Separation is stated from the base of the collimating nozzle

jammed. For this reason a new lower plate was produced out of black polypropylene. Calculations were made to ensure that the four horizontal screws binding into this material would be strong enough to support the weight of the whole can full of water.

4.4.3 Pre-Deployment Tests

Beam Width and Structure in Air

The actual collimating length of the nozzle is 34.9 mm. With a 4 mm inner diameter, the half-angle of the beam should be 6.6° for perfect operation. This requires the source to be vertical, with the two steel disks in good alignment. The diameter of the beam image was measured using a 60 W light bulb inside the mask with a disk of white tissue paper to diffuse the source of light. The image was created on photographic paper for different path lengths in air. The measured widths of these images are given in table 4.3.

Figure 4.7 shows the measurements from a densitometry scan across the image of the beam obtained at 85 cm separation in both horizontal and vertical directions. The beam appeared sharply collimated and the beam width measured was used as input for the first Monte Carlo simulations of masked laserball data.

Drainage

Approximately one litre of air is contained in the volume between the laserball and the mask, which is replaced by D_2O as the source is deployed in the detector. Water can enter through the hole in the collimator and air can escape through "bleeding holes" in the upper plate and shield of the mask. The reverse occurs when the source is removed from the detector.

The drainage rate was measured using a spare manipulator motor system at Queen's University. The tension in the manipulator rope increased steadily over a period of 25 minutes before reaching a maximum of 36.8 N. Based on this, a 30 minute pause was required after deployment in the D_2O to ensure complete evacuation of air from the cham-

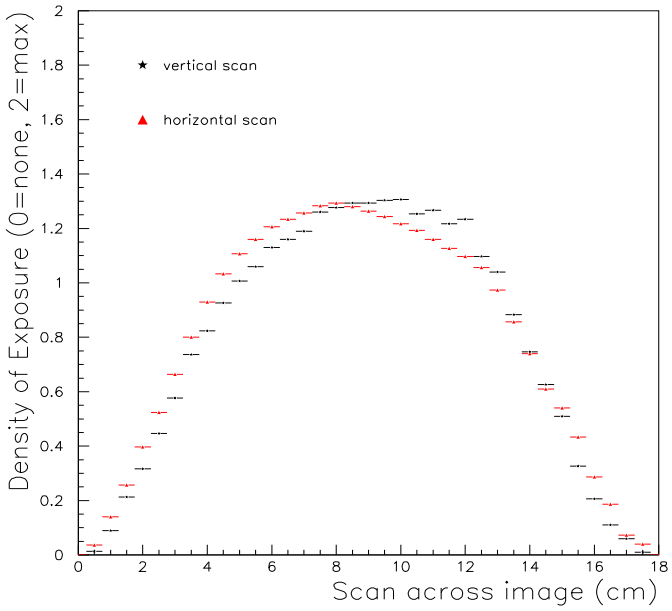


Figure 4.7: Results of a densitometry scan of the image of the collimated beam at 85 cm separation.

ber. Complete drainage of water from the mask volume took 1.5 hours after removal from the water.

4.4.4 Cleanliness

The mask was constructed from three materials, all of which were tested for impurities that could leach into the D_2O. The bulk of the mask, including all screws, was made of low-carbon stainless-steel, which has been previously tested and approved for use in many of the calibration sources used in SNO. Black polypropylene was used only for the collimator and lower sliding plate, as tests indicated that low levels of radium may leach out of this material. The shield was made of white polypropylene, as reflections from this part of the mask were less critical and leach-test results were more favourable for this material.

All pieces of the source were cleaned ultrasonically and handled only when wearing gloves. To avoid the introduction of Radon gas into the detector, the mask was flushed with nitrogen gas prior to deployment.

4.5 Monte Carlo Simulation of the Masked Laserball

The analysis was carried out on both data and simulated events for each run. Intensity fluctuations of the laserball were not modelled, and a consistent light level of 65 photons per laser pulse was assumed. The laser ball geometry was included in the simulation but the mask geometry was not.[¶] The simulated pulses of photons were emitted into a cone of half-angle 7.25° with the measured timing distribution of the laserball source.[‖] The angular distribution of light within the cone was modelled with a Gaussian distribution of $\sigma = 7.25°$. The optical parameters used for these simulations were based on preliminary optical fits to laserball data, which best represented the understanding of the detector at the time of analysis.

4.6 Comparisons of Data and Simulations

Table 4.4 contains information on the masked laserball data obtained for this analysis in April 2002. The main scattering analysis was carried out using the "z position scans". A series of runs with different laserball intensity were taken for systematic and noise studies, whilst a horizontal scan was carried out to to check against unusual geometric effects.

Comparisons were made of the timing and angular distribution of the data and simulated events. The left hand plot of figure 4.8 shows the data and Monte Carlo timing distributions for run 24438 (λ =337 nm, z =500 cm). The general timing structure of the beam is replicated by the simulation but small differences are expected due to the effects of the mask geometry, which was omitted from the simulation.

The right hand plot of figure 4.8 shows the data and Monte Carlo angular distributions for the same run. The angle is taken relative to the $-z$ axis from the source position. Agreement of the angular distribution in the beam-line is not perfect, indicating that there is some leakage of light at the edges of the beam. However, the level of light immediately outside the "theoretical beam" is two orders of magnitude below the beam intensity and scattering of this leaked light should have a negligible effect on the number of hits observed in the analysis region. As explained in the following section, all PMT hits were used to determine the amount of light deposited in the detector, and the analysis criteria were defined to exclude the region close to the beam edges where data-simulation discrepancies were observed.

[¶]Inclusion of the mask geometry would improve the accuracy of simulations, but is beyond the scope of this thesis.

[‖]The angle was based on the pre-deployment tests and preliminary data.

Run Number	Source position			λ(nm)	ND	t (minutes)
	x(cm)	y(cm)	z(cm)			
z position scans						
24438	-2.6	-10.3	506	337	1.5	15
24440	-0.14	-1.0	309.6	337	1.5	15
24442	0.0	-1.0	9.87	337	1.6	20
24473	-2.5	-10.3	506.2	365	1.5	15
24478	-2.5	-10.3	506.2	421	1.2	15
24480	-2.5	-10.3	506.2	365	1.5	15
24486	-0.1	-0.5	309.8	500	1.5	15
24488	-0.1	-0.5	309.8	421	1.2	15
24492	-0.1	-0.5	309.8	365	1.5	15
24494	-0.1	-0.9	9.8	365	1.6	15
24496	-0.1	-0.9	9.8	365	1.6	5
24501	-0.1	-0.9	9.8	421	1.2	20
24503	-0.1	-0.9	9.8	500	1.6	20
Intensity scan						
24432	-2.6	-10.3	505.8	365	1.6	15
24434	-2.6	-10.3	505.8	365	1.2	15
24436	-2.6	-10.3	505.8	365	1.0	15
x position scans						
24444	100	0	10	337	1.2	10
24446	210	0	10	337	1.0	10
24448	300	0	10	337	1.0	10
24450	387	0	10	337	1.0	10
24452	457	0	9	337	1.0	10
24454	486	0	12	337	1.0	10
24456	486	0	12	365	1.5	10
24458	458	-1	10	365	1.5	10

Table 4.4: A summary of data taken with the laserball and mask used for this analysis. The table gives the run number, the laserball source position in detector coordinates (the point of collimation is 10 cm directly below the laserball centre), the wavelength, λ, the value of neutral density filter used, ND, and the run duration, t, in minutes.

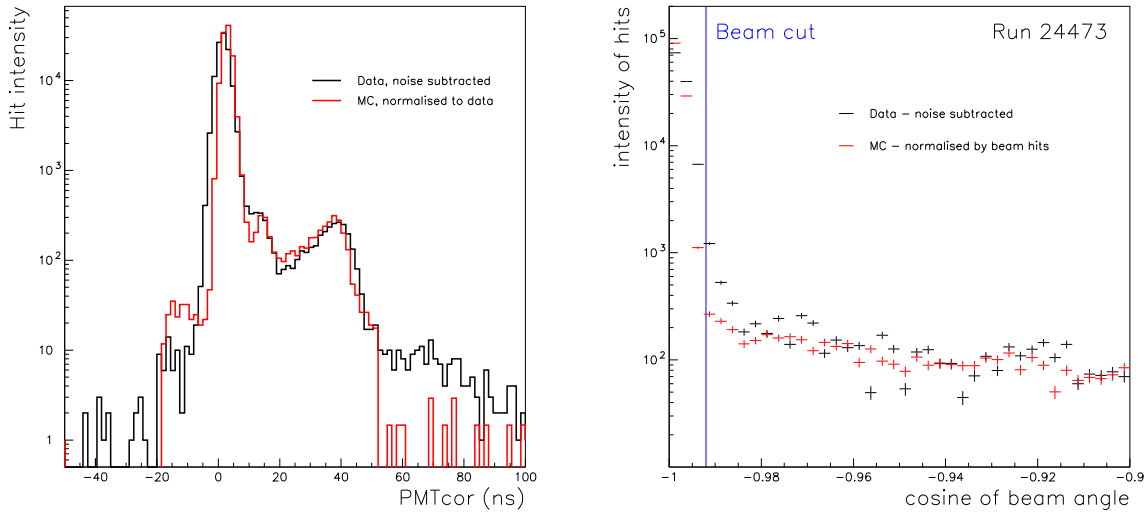


Figure 4.8: Comparison of the beam timing (left) and the total angular distribution (right) for data and simulations of run 24438.

4.7 Analysis Procedure

The basic analysis involved counting up the number of hits, N_{a}, in a selected analysis region over the entire run. This was divided by the total number of hits, N_{total}, to obtain the fraction of light scattered into that region, F_{scat}.

$$F_{\mathrm{scat}} = \frac{N_{\mathrm{a}}}{N_{\mathrm{total}}} \tag{4.3}$$

The same analysis was applied to both data and Monte Carlo simulated events and the two fractions were compared. $R(F_{\mathrm{scat}})$ thus gives the factor by which the simulated scattering should be scaled to reproduce the data.

$$R(F_{\mathrm{scat}}) = \frac{F_{\mathrm{scat}}(\mathrm{data})}{F_{\mathrm{scat}}(\mathrm{MC})} \tag{4.4}$$

Only data events triggered by the EXTA trigger** were considered, and no $Nhits$ threshold was applied.

4.7.1 A Note on Timing Distributions

PMT firing times were one of the main observables used in this analysis. By correcting the firing time for different flight paths, different features of the data could be selected, as illustrated schematically in figure 4.9.

**The External Asynchronous (EXTA) trigger is received directly from the laser, and is not linked to either the 10 MHz or 50 MHz clock.

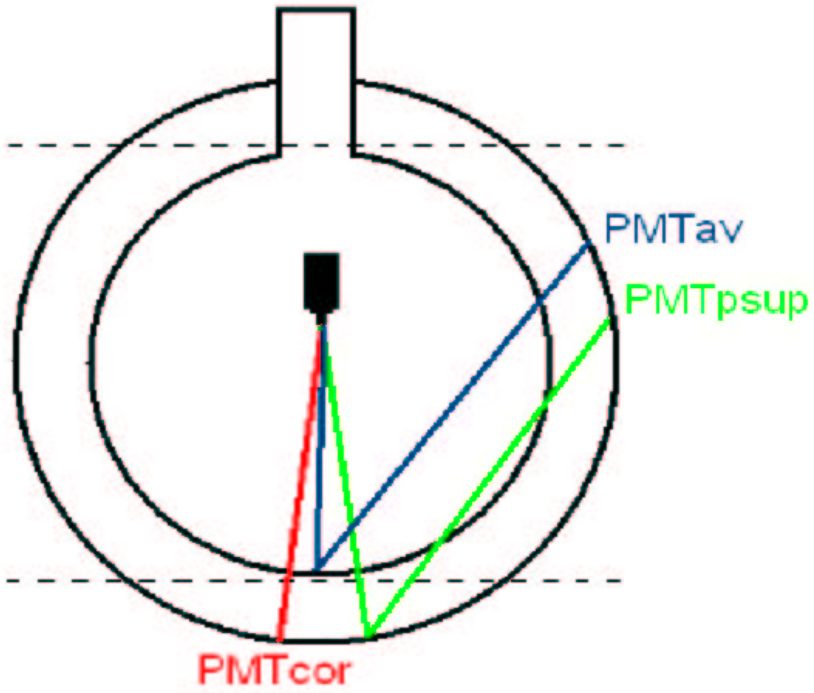


Figure 4.9: A schematic cross-section of the SNO detector with lines to illustrate the different light paths selected by the timing distributions PMTcor (direct light - equation 4.5), PMTpsup (reflections from the PSUP - equation 4.6) and PMTav (reflections from the AV - equation 4.7). The dashed horizontal lines indicate the spatial constraints placed on the analysis region. This diagram is not to scale.

The first timing distribution used, *PMTcor*, was a simple correction for the time of flight from the source position to the PMT:

$$PMTcor = PMTT - \frac{\vec{SP}}{c} - O_{\mathrm{t}} \tag{4.5}$$

where *PMTT* is the PMT firing time, $\vec{SP}$ is the distance between the source and the tube hit, c is the speed of light in D_2O at that wavelength, and O_{t} is the timing offset of the laser pulse in the trigger window. O_{t} was fitted to the peak in the beam timing distribution separately for data and Monte Carlo for each run.

PMTpsup was defined to select light reflected off the PSUP:

$$PMTpsup = PMTT - \frac{\vec{SB} + \vec{BP}}{c} - O_{\mathrm{t}} \tag{4.6}$$

where $\vec{SB}$ is the path length from the source to the PSUP directly below, at $z = -850\,\mathrm{cm}$, and $\vec{BP}$ is the path length from that point on the PSUP to the PMT that was hit.

Finally, *PMTav* was defined to select light that had reflected, or scattered, off the AV:

$$PMTav = PMTT - \frac{\vec{SA} + \vec{AP}}{c} - O_{\mathrm{t}}. \tag{4.7}$$

This is similar to the definition for *PMTpsup*, but with a path length defined by the vectors $\vec{SA}$, the path from the source to the acrylic vessel at $z = -600\,\mathrm{cm}$, and $\vec{AP}$, the path from the AV at that point to the PMT that was hit.

Figure 4.10 shows the timing distribution of data and simulated events for a single masked laserball run. Noise hits, which have not been simulated, can be clearly seen at early firing times, as well as contributing across the time window. The sharp, central peak is due to prompt light in the beamline. The two smaller peaks between 225 ns and 250 ns, seen most clearly in the zero scattering distribution, are due to reflections off the PSUP and AV. Comparison of the simulated distributions with and without Rayleigh scattering shows that scattering hits should appear across a wide range of the timing distribution.

The PMT timing distribution for any given PMT has an intrinsic shape which has been extensively studied with isotropic laserball data [81]. Whilst the majority of incident photons produce a prompt electronic signal at the PMT, there are mechanisms which result in a small fraction arriving earlier or later in time to produce three identifiable peaks. Firstly the phenomenon of "pre-pulsing" can lead to a PMT signal 15 ns before the prompt peak. This occurs when an incident photon passes through the PMT envelope and photocathode without producing a photoelectron and strikes the first dynode directly. Secondly, "late-pulsing" occurs if photons escape from the first dynode and return to the photo-cathode, sending another photoelectron down the dynode stack and resulting in

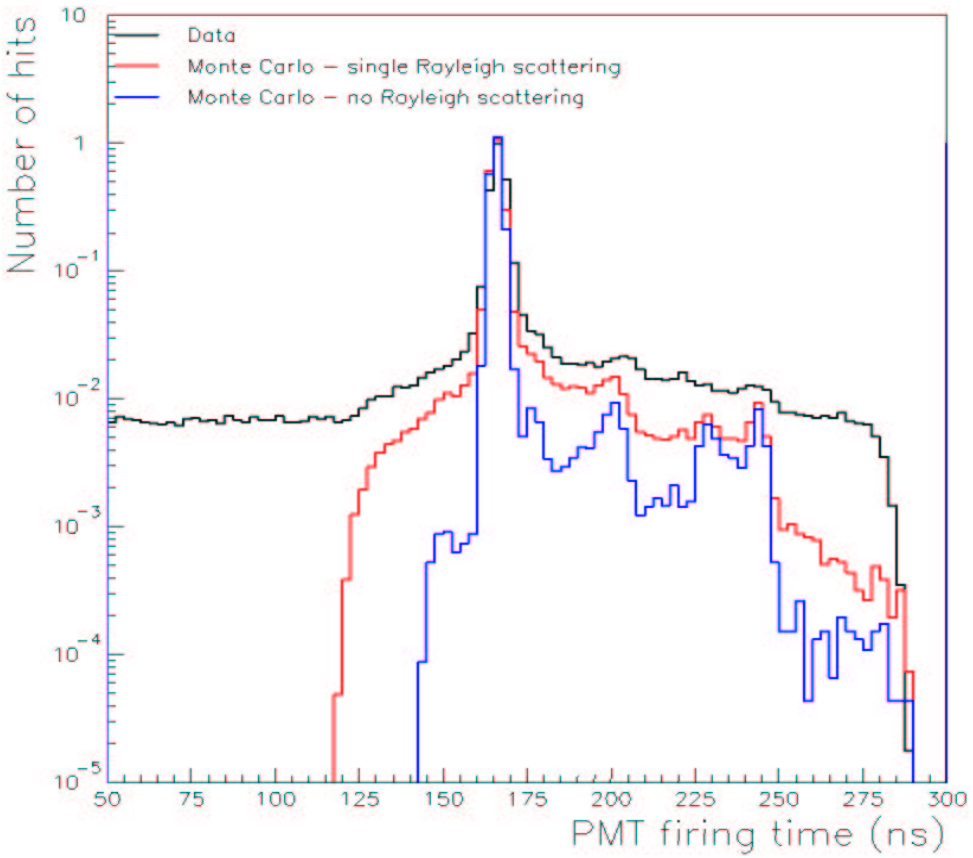


Figure 4.10: The distribution of PMT hits plotted against PMT firing time for data (black), simulated events with theoretical Rayleigh scattering (red) and with zero Rayleigh scattering (blue). Noise hits were not included in the simulation.

a signal delayed by 15 ns. Finally, "after-pulsing" is thought to be due to collisions of photo-electrons on residual gas molecules in the PMT. The ionized gas molecules then drift slowly up to the cathode, resulting in a delayed PMT signal. The diffuse timing distribution of these after-pulses is dependent on the mass of the gas molecules and can extend to μs after the prompt peak.

4.7.2 Noise

The PulseGT was used to determine the total number of noise hits in each run, $N_{\rm nt}$, and also the number of noise hits falling inside analysis cuts, $N_{\rm na}$. In each case the value was normalised to the number of EXTA triggered events in the run.[††]

The noise hits at early times in the trigger window ($PMTcor \leq 0$ ns) were compared between runs to test for changes in noise levels associated with source intensity, position or wavelength. The timing distributions for the intensity run scans are shown in figure 4.11. Despite the change in intensity, there is no evidence for a change in noise levels, as seen by the agreement of the three distributions in the region $-150 < PMTcor < -20$ ns in this figure.

4.7.3 Total Hits

To determine the total amount of observable light[‡‡] deposited in the detector, it was necessary to account for the effect of multiple hits which was introduced in section 3.4. Due to the high photon intensities, there is a higher probability for PMTs in the beamline to be struck by more that one photon, even though only one hit is ever registered. Whilst the charge measured on a given PMT is generally higher in the event of multiple photo-electrons, the charge resolution is relatively poor and therefore charge information was not used for analysis purposes.

The expected number of hits missed due to multiple hit effects, $N_{\rm mh}$, for PMTs in the beamline was calculated using equation 3.11. The value of $\mu(i)$ was obtained from:

$$\mu(i) = \frac{Nhits(i) + N_{\rm missed}(i)}{N_{\rm events}} \approx \frac{Nhits(i)}{N_{\rm events}} \tag{4.8}$$

where $Nhits(i)$ is the number of hits registered by PMT i over the entire run, $N_{\rm missed}(i)$ is the number of secondary photons which go undetected, and $N_{\rm events}$ is the number of events in the whole run. $N_{\rm missed}(i)$ is unknown but a leading order approximation was made for $\mu(i)$ which was applicable for the low intensity regime considered here.

[††]This was approximately a factor of nine as the laserball was pulsed at a rate of 45Hz whilst the PulseGT has a rate of 5Hz. However, exact event numbers were used to account for fluctuations in the laser pulsing rate.

[‡‡]The total light convolved with the detector efficiency.

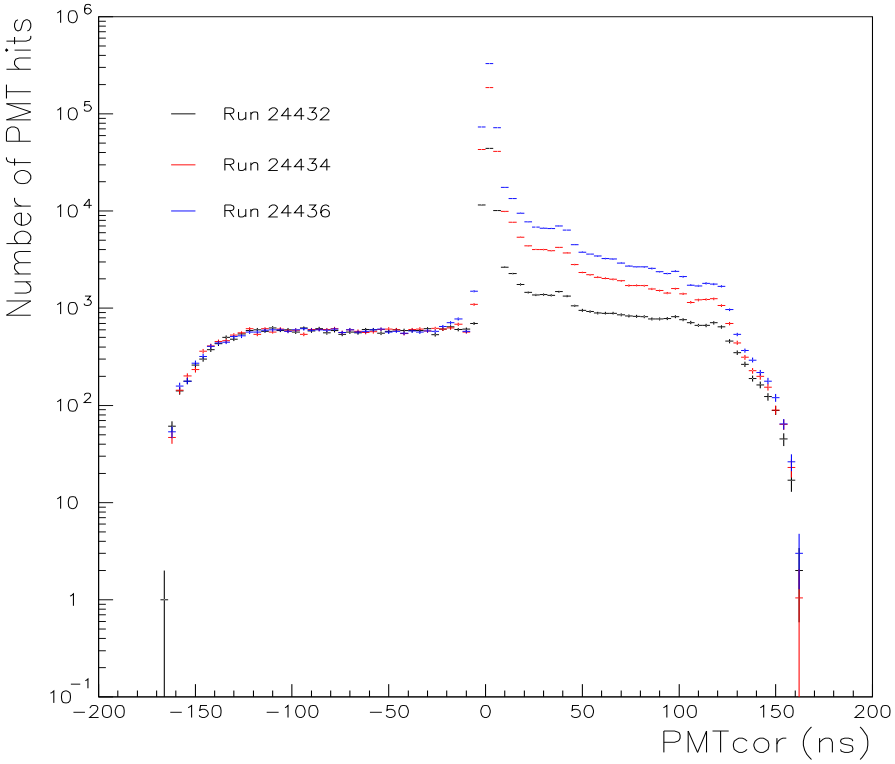


Figure 4.11: Data timing distribution of three different intensity runs. Despite the change in laserball intensity there is no evidence for a change in the noise levels, as seen by the agreement of the three distributions in the region $-150 < PMTcor < -20$ ns.

Equation 4.9 shows how these additional hits were combined with the number of hits in the beamline, N_{beam}, the number of hits in the rest of the detector, N_{out}, and also the total noise hits for the run to obtain the total amount of light:

$$N_{\mathrm{total}} = N_{\mathrm{beam}} + \left(\sum_{i=1}^{N_{\mathrm{bP}}} \frac{N_{\mathrm{mh}}(i)}{Nhits(i)} \times N_{\mathrm{events}} \right) + N_{\mathrm{out}} - N_{\mathrm{nt}} \tag{4.9}$$

where N_{bP} is the number of PMTs in the beamline. The multiple hit correction, which required separate treatment of each PMT, was only applied to PMTs in the beam line. Outside the beamline, the correction was negligible since each PMT was typically hit in less than 0.02% of triggers, with no PMT contributing to more than 1% of the triggers in a typical run. In the beamline, the average PMT contributed to 2.9% of triggers, whilst the most frequently hit PMT was involved in 12.5% of the triggers. The uncertainty in the multiple hit correction was conservatively estimated at 5%, based on the spread of the correction values for different runs.

4.7.4 The Analysis Region

An analysis region, free from prompt beamline hits and dominant reflections, was defined by spatial and timing criteria. The number of hits in this region, N_{ha}, also included a number of hits due to PMT dark noise, N_{na}, which was subtracted to obtain the number of hits due to scattering, N_{a}:

$$N_{\mathrm{a}} = N_{\mathrm{ha}} - N_{\mathrm{na}} \tag{4.10}$$

The spatial selection criterion was defined in detector coordinates such that only PMTs hits in the region $-0.75 < \cos(\theta) < 0.75$ were selected, where θ is the vertical angle of the PMT with respect to the centre of the detector. These cuts were chosen to minimise contamination by leakage of light around the beam edges. Also, differences between data and simulation were not well understood at high $\cos(\theta)$ as the source geometry was not simulated and the neck region of the detector has more complex geometry.

The timing selection criterion was defined in terms of the flight time for light reflected off the AV, *PMTav*, as given in equation 4.7. Only hits with firing times satisfying $PMTav < -5\,\mathrm{ns}$, which arrive before reflections off the acrylic, were included in the scattered data set.

These criteria are illustrated in figure 4.12, in which each point represents a single PMT hit. The blue points (at the bottom) indicate hits on PMTs falling directly in the beam line, whilst red indicates hits that satisfy the analysis criteria. The green lines show the spatial cuts to select the analysis region. The contribution of scattered photons is

seen in the dark red band starting at $PMTcor = 0$ ns. To the right of the analysis region, two diagonal features due to reflection off the AV and PSUP can be seen.

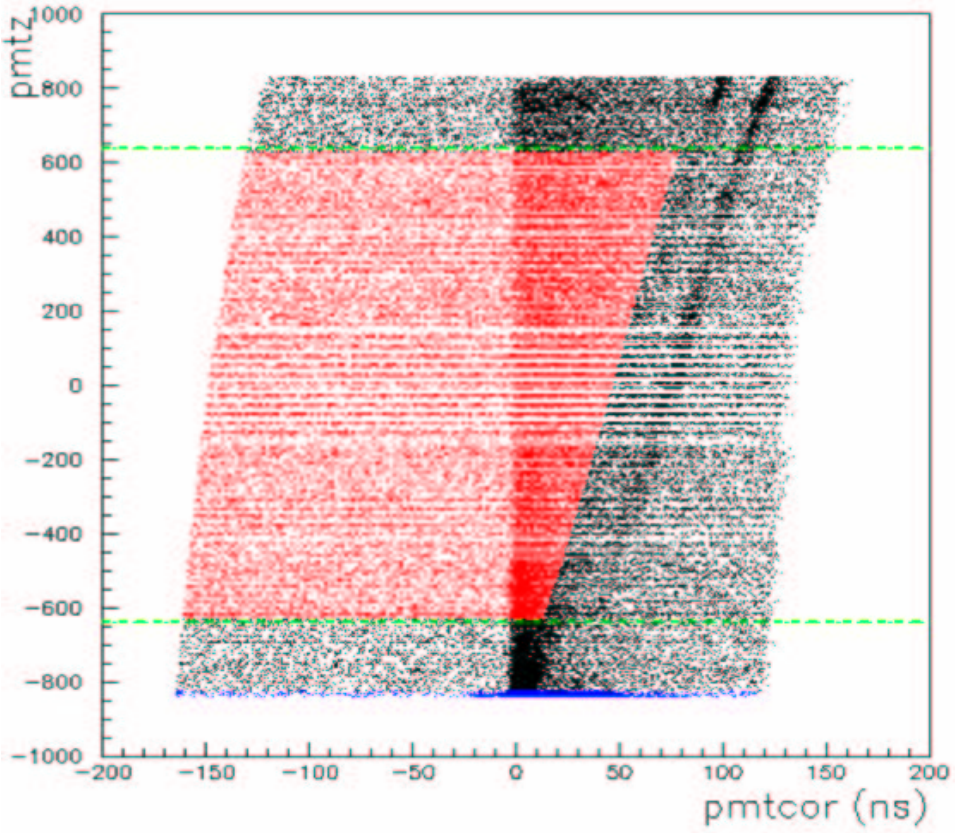


Figure 4.12: Time ($PMTcor$ as defined in equation 4.5) and spatial coordinates of PMT hits in a masked laserball run. $PMTz$ is the vertical position of the PMT relative to the centre of the AV.

4.8 Results

The extracted values of $R(F_{\text{scat}})$ are given in table 4.5 and plotted against wavelength and source z position in figure 4.13. The weighted mean over all source positions and wavelengths is

$$\langle R(F_{\text{scat}}) \rangle = 1.303 \pm 0.027. \tag{4.11}$$

This value is larger than unity, indicating that there is slightly more scattering than predicted by simulations of pure Rayleigh scattering. However, the magnitude of this additional scattering is small, suggesting that the level of particulate contamination in the D_2O is very low.

Run	Source z Position (cm)	$R(F_{\text{scat}})$
$\lambda = 337\,\text{nm}$		
24438	500	1.312 ± 0.054
24440	300	1.159 ± 0.062
24442	0	1.309 ± 0.077
$\lambda = 365\,\text{nm}$		
24473	500	1.142 ± 0.073
24492	300	1.168 ± 0.135
24494/96	0	1.293 ± 0.259
$\lambda = 421\,\text{nm}$		
24478	500	1.589 ± 0.089
24488	300	1.532 ± 0.096
24501	0	1.561 ± 0.134
$\lambda = 500\,\text{nm}$		
24480	500	1.668 ± 0.246
24486	300	1.191 ± 0.134
24503	0	0.355 ± 0.438

Table 4.5: $R(F_{\text{scat}})$ for each masked laserball run. (Approximate source positions are given here - see table 4.4 for the precise locations.

There is no strong evidence for a trend in $R(F_{\text{scat}})$ with either source position or wavelength so the approximation to a single Rayleigh-like scattering length appears reasonable. If the scattering mechanism occurred with a grossly different angular distribution, the ratio of scattering seen in the analysis region would change with source position. Equally, a change with wavelength, λ, would indicate a deviation from the expected λ^{-4} dependence of Rayleigh scattering

4.9 Sensitivity to H_2O Scattering

The analysis discussed in this chapter was developed to be sensitive to scattering in the D_2O volume, but was expected to be relatively insensitive to scattering in the H_2O due to the spatial selection criterion. The path length through the H_2O is only 2.5 m and a high proportion of light scattered in this region region will hit PMTs close to the beamline rather than in the analysis region.

To check this hypothesis, the analysis was repeated with two different simulated

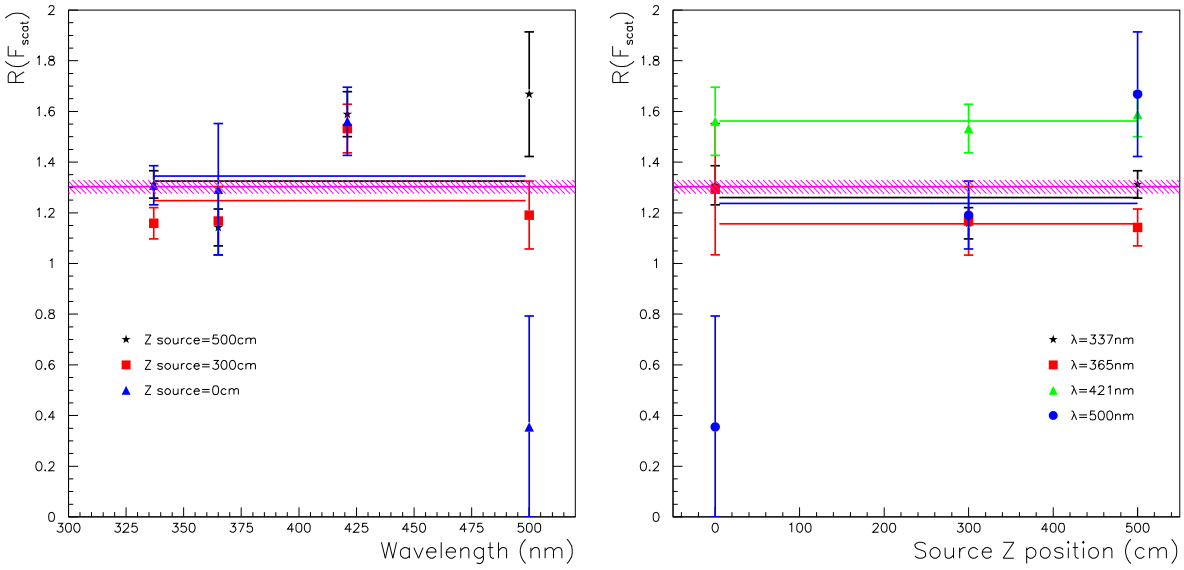


Figure 4.13: The ratio of data to Monte Carlo scattering fractions, $R(F_{\text{scat}})$ for each run plotted against wavelength on the left, and source z position on the right. The weighted mean with associated error is given by the pink hashed band.

data sets. In one set, both the D_2O and H_2O Rayleigh scattering lengths were scaled by a factor of 1.3 whilst, in the second set, only the D_2O scattering length was modified. The mean values of $R(F_{\text{scat}})$ were in good agreement for the two different Monte Carlo scenarios, indicating that the analysis is, indeed, insensitive to H_2O scattering.

4.10 Verification of the Analysis Technique

To verify the sensitivity of this analysis to scattered light, the calculation was extended to extract the Rayleigh-like scattering length from the Monte Carlo data sets. To do this a number of approximations were made. For example, the contribution of attenuation in the H_2O and AV was assumed to be negligible.

The first step of this calculation was to estimate the amount of scattered light excluded from the analysis region.

4.10.1 Hits Excluded by the Timing Criterion

To correct for the amount of scattered light excluded by the timing criterion, a second simulated data set was created for each run with no contribution from Rayleigh scattering. For this second simulation, the hits excluded by the timing criteria were assumed to all

Source z position (cm)	F_{sample}
500	0.5641
300	0.5639
0	0.5338

Table 4.6: The calculated fraction of the Rayleigh distribution sampled inside analysis cuts for different source positions along the central axis of the detector.

be due to reflections. The number of scattered hits omitted from the analysis region was estimated by subtracting the hits failing the timing criterion in this second simulation, $MC_B(N_{\text{cut}})$, from those failing the same criterion for the original simulation, $MC_A(N_{\text{cut}})$.

The two simulations were first normalised by the total number of hits, N_{tot}, as calculated with equation 4.9. The number of simulated scattering hits excluded from the analysis region was:

$$N_{\text{scatlost}} = MC_A(N_{\text{cut}}) - \left(\frac{MC_B(N_{\text{cut}})}{MC_B(N_{\text{tot}})}\right) \times MC_A(N_{\text{tot}}). \tag{4.12}$$

4.10.2 Fraction of Rayleigh Distribution Sampled

The applied spatial criterion selected only a fraction of the total scattered light. For each source position, a numerical integration of the theoretical Rayleigh distribution (equation. 4.13) was carried out over the path length through the D_2O to obtain the fraction of the distribution sampled.

$$F_{\text{sample}} = \frac{3}{8}\int_{\cos(\theta_{\text{min}})}^{\cos(\theta_{\text{max}})} \left(1+\cos^2\theta\right) d\cos\theta \tag{4.13}$$

The fractions evaluated for each source position are summarised in table 4.6. The uncertainty on this value was estimated to be $\pm 5\%$.

4.10.3 Scattering Length Calculation

The scattering length was determined from the corrected fraction of light scattered, F_{S}:

$$F_{\text{S}} = \frac{N_{\text{a}} + N_{\text{scatlost}}}{N_{\text{total}} \times F_{\text{sample}}} \tag{4.14}$$

The scattering length, Λ_{S}, is related to the fraction of scattering, F_{S} by:

$$F_{\text{S}} = 1 - e^{-\left(\frac{x}{\Lambda_{\text{S}}}\right)} \tag{4.15}$$

where x is the photon path length through the D_2O.

Wavelength (nm)	Rayleigh Scattering Length (m)	
	Extracted	Expected
337	68.3 ± 4.8	75.3
365	98.8 ± 6.9	106.8
421	176.5 ± 12.5	197.5
500	420.6 ± 31.8	408.8

Table 4.7: Comparison of the extracted and expected Scattering lengths for Monte Carlo data sets.

Wavelength (nm)	New Rayleigh Scattering Length (m)
337	58.72
365	83.20
421	154.00
500	318.62

Table 4.8: New Rayleigh scattering lengths as used in the salt optics simulation.

Table 4.7 gives the scattering lengths extracted from the Monte Carlo data sets. In all but the 500 nm case, the extracted lengths were shorter than those expected, which is probably due to optical effects in the AV and H_2O regions that were not considered in this analysis. The dominant effect is the absorption of light in AV which increases exponentially at short wavelengths. (At 500 nm the absorption coefficient is $< 1\%$ but at 300 nm it is $\approx 50\%$.) Nonetheless, the scattering lengths are of the correct order of magnitude, indicating that the analysis is sensitive to the scattering phenomena expected.

4.11 Improved Optical Model

The results measured with the masked laserball were used to update the optical model of the SNO detector. The magnitude of Rayleigh scattering was increased by the factor $\langle R(F_{\rm scat})\rangle$ and the contribution of this scattering was subtracted from the attenuation determined by the optical fit to obtain the magnitude of absorption in the D_2O. The scattering lengths determined (in SNOMAN) by scaling the amount of Rayleigh scattering in the D_2O by a factor of 1.289* are given in table 4.8.

*This scale factor was obtained from an earlier analysis of the masked data. A small error was later found in the analysis and has been corrected in the results presented here. However, as the effect of this error on $R(F_{\rm scat})$ was small ($< 1\sigma$) the value used to update the model was not changed.

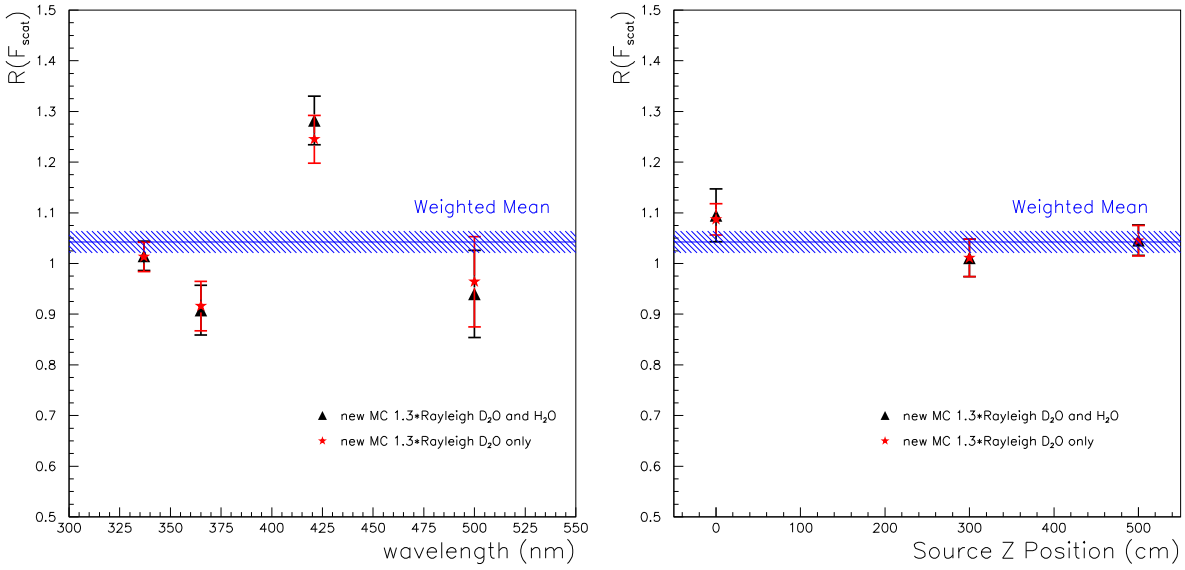


Figure 4.14: The ratio of scattering fractions, $R(F_{\rm scat})$ between data and events simulated with Rayleigh scattering increased by a factor of 1.289. $R(F_{\rm scat})$ is plotted against wavelength on the left, and source z position on the right. The weighted mean with associated error is given by the blue hashed band.

The analysis was repeated, comparing the scattering fraction obtained from data to that obtained from simulations with the updated scattering levels. The values of $R(F_{\rm scat})$ obtained are presented in figure 4.14. In the left-hand plot, the weighted mean over the three source positions is plotted against wavelength, whilst the weighted mean over the four wavelengths is plotted against source position in the right hand plot. The weighted mean over all positions and wavelengths is slightly greater than one, as shown by the hashed band in these plots.

4.12 Summary

In this chapter, the first direct measurement of large-angle, Rayleigh-like scattering in the SNO D_2O volume has been presented. The equivalent scattering observed is a factor of 1.3 greater than the calculated amount of nominal Rayleigh scattering. Therefore, the level of particulate contamination in the D_2O must be very low.

Detector simulations have been corrected by scaling the amount of Rayleigh-like scattering in the model. This improved knowledge of scattering phenomena in the D_2O helps to limit the magnitude of possible systematic differences between observed and simulated observable parameters, which are used to separate events due to different neutrino

interactions in the salt data set as discussed in the next chapter.

Data obtained using the masked laserball may be used to study other optical phenomena in the SNO detector and to improve on detector simulations. For example, the directional source allows better separation of reflections off the AV and PSUP.

A directional source may also prove extremely useful for optical calibration of the SNO detector in the third phase of the experiment when the presence of neutral-current detectors in the D_2O volume will cause shadowing effects that must also be accurately modelled.

Chapter 5

Signal Extraction Using Maximum Likelihood

The maximum likelihood technique is the universally most powerful method of parameter estimation. It endeavours to find the "most likely" values of the parameters that describe a set of data, by maximising a product of probabilities known as a *likelihood function*. The likelihood function is parameterised to make efficient use of the information available and is ideally suited to the problem of separating the different event classes detected by the SNO experiment. This chapter introduces the maximum likelihood technique and its application to the SNO data analysis, along with the analysis tools developed for this thesis project.

5.1 Maximum Likelihood Technique

The maximum likelihood technique requires a data set whose individual components have distinct *observables* that can be used to distinguish them. An example of such an observable is $\cos\theta_\odot$, the cosine of the angle between the fitted event direction and the direction to the Sun. The distribution in $\cos\theta_\odot$ is strongly peaked at 1.0 for the ES signal, shows a $(1 - \frac{1}{3}\cos\theta_\odot)$ distribution for the CC signal and is flat for the NC signal. The overall distribution of $\cos\theta_\odot$ for the data can thus be fit to a linear superposition of the three individual signal distributions. The weighting factors for each signal in the best fit yield the best estimate for the relative proportions of each component of the data sample.

The likelihood function calculated for each event is $\mathcal{L}$:

$$\mathcal{L} = \sum_{i=1}^{N_s} (P(i/j_1, ...j_{N_{ob}})) \tag{5.1}$$

where $P(i/j_1, ...j_{N_{ob}})$ is the probability that an event is of signal type i given the data values for the set of N_{ob} observable parameters, j and N_s is the number of signals fitted

for. For the case $N_{\rm ob} = 1$, this probability is given by equation 5.2:

$$P(i/j) = \frac{n_i}{N_{\rm data}} \times F(i/j) \tag{5.2}$$

where $F(i/j)$ is a Probability Density Function (PDF), that describes the probability for a given variable, j, to be drawn from a particular event class i. $N_{\rm data}$ is the total number of events in the data set and n_i is the number of events of class i. The set of n_i values are varied to maximise the likelihood function. $N_{\rm data}$ adopts a Poisson distribution about the true mean, μ_0. Thus, the probability of obtaining $N_{\rm data}$ events should also be included in the total likelihood, L_t, which is calculated from the product of the individual event likelihoods, $\mathcal{L}_d$:

$$L_t = \frac{\mu_0^{N_{\rm data}}}{N_{\rm data}!} e^{-\mu_0} \prod_{d=1}^{N_{\rm data}} \mathcal{L}_d(n_i, ..n_{N_{\rm s}}) = \frac{e^{-\mu_0}}{N_{\rm data}!} \prod_{d=1}^{N_{\rm data}} \mu_0 \mathcal{L}_d(n_i, ..n_{N_{\rm s}}) \tag{5.3}$$

This is known as the *extended likelihood function.*

Due to the large values obtained for L_t it is usually more precise and convenient to compute and maximise the log likelihood, $\mathbb{L}$, which is obtained by summing the logarithm of the likelihood over all data events:

$$\mathbb{L} = -\mu_0 - \log(N_{\rm data}!) + \sum_{d=1}^{N_{\rm data}} \log(\mu_0 \mathcal{L}_d(n_i, ..n_{N_{\rm s}})) \tag{5.4}$$

The expected number of events of type i can be defined as:

$$\mu_i = \mu_0 \frac{n_i}{N_{\rm data}} \tag{5.5}$$

such that

$$\mu_0 = \sum_{i=1}^{N_{\rm s}} \mu_i \tag{5.6}$$

As μ_0 and $N_{\rm data}$ are both constant, the second term in 5.4 can be omitted from the function to be maximised, and the number of events, n_i can be used instead of μ_i. Therefore, the log likelihood can be written:

$$\mathbb{L} = \sum_{d=1}^{N_{\rm data}} \log \left(\sum_{i=1}^{N_{\rm s}} n_i F(i/j_1(d), ...j_{N_{\rm ob}}(d)) \right) - \sum_{i=1}^{N_{\rm s}} n_i \tag{5.7}$$

Maximisation of this function provides the best estimate for the set of parameters n_i that describe the data set.

5.1.1 Constraining Parameters in the Fit

The likelihood function can be modified to incorporate additional constraints on the fit which might be obtained by independent means. For example, calibration source data may determine that a certain background will contribute $n_i = X^{+a}_{-b}$ events, where a and b are one sigma errors. This additional information can be included in the fit through the probability C_i, based on an asymmetric Gaussian with mean X.

$$\mathbb{L} = \sum_{d=1}^{N_{\rm data}} \log\left(\sum_{i=1}^{N_{\rm s}} n_i F(i/j_1(d), ...j_{N_{\rm ob}}(d))\right) - \sum_{i=1}^{N_{\rm s}} n_i + \sum_{i=1}^{N_{\rm s}} \log(C_i) \tag{5.8}$$

When n_i deviates by one sigma from X, $\mathbb{L}$ will decrease by 0.5. As shown, in section 5.2, $(\mathbb{L}_{\rm max} - 0.5)$ is the value of the log-likelihood for the 1σ confidence interval. For unconstrained fit parameters, C_i is set to unity.

5.1.2 Incorporating Additional Parameters in the Fit

In addition to the flux of each signal type, there may be other parameters of interest in the likelihood fit, such as systematic variations in the observable parameters. For example, suppose that the mean value of observable j is not precisely known so that the nominal PDF in j may not match the data correctly. A shift in the mean could be included as an additional fit parameter and the additional systematic uncertainty involved contributes to the errors returned by the fit through the extra degree of freedom allowed. This approach will naturally deal with the correlations between multiple systematic uncertainties. Chapter 7 describes the inclusion of systematic uncertainties by this technique in an energy-constrained fit to the salt data set.

5.2 Statistical Uncertainties on Fit Parameters

A number of methods can be used to determine the uncertainties on parameters fit with the Maximum Likelihood technique. The most "brute force" approach is a Monte Carlo method, which involves performing the fit on a large number of simulated experiments. The distribution of values for each fit parameter can, in principle, then be used to determine the error. However, this method is computationally slow and could be complicated by systematic differences between the real data set and the Monte Carlo simulations. This technique has been used when carrying out bias tests, to verify errors obtained by other methods.

The *Rao-Cramer-Frechet* (RCF) inequality [82], which can be determined analytically, gives a lower bound on the variance of a given likelihood fit parameter, f. For a

sufficiently large data sample, the variance can be estimated from the second derivative of the log-likelihood function, $\mathbb{L}$, with respect to f, as given in equation 5.9:

$$\hat{\sigma}^2 = \left(\frac{-1}{\frac{\delta^2 \mathbb{L}}{\delta f^2}} \right) \Bigg|_{f=\hat{f}} \tag{5.9}$$

This technique is employed by the MINUIT [83] package, which numerically determines the matrix of second derivatives for the best fit point and inverts it to find the covariance matrix.

A third technique is a graphical method, which has been added to the analysis code used in this thesis and provides an excellent check of errors returned by MINUIT. As a simple example, consider the log-likelihood function, $\mathbb{L}$, for a single fit parameter, f. Expanding that function as a Taylor series about f, gives equation 5.10.

$$\mathbb{L}(f) = \mathbb{L}(f) + \left[\frac{\delta \mathbb{L}}{\delta f} \right]_{f=\hat{f}} (f - \hat{f}) + \frac{1}{2!} \left[\frac{\delta^2 \mathbb{L}}{\delta f^2} \right]_{f=\hat{f}} (f - \hat{f})^2 + ... \tag{5.10}$$

However, we know that when f takes the best fit value $\hat{f}$, $\mathbb{L}(f)$ takes its maximum value, $\mathbb{L}_{\max}$, so the first differential is zero. Ignoring higher order terms, equation 5.10 reduces to equation 5.11:

$$\mathbb{L}(f) = \mathbb{L}_{\max}(f) - \frac{(f - \hat{f})^2}{2\sigma_{\hat{f}}^2} \tag{5.11}$$

Setting f to $\hat{f} + \sigma_{\hat{f}}$ yields

$$\mathbb{L}(\hat{f} + \sigma_{\hat{f}}) = \mathbb{L}_{\max}(f) - \frac{1}{2}. \tag{5.12}$$

Errors can, thus, be determined graphically on parameter f by stepping away from the best fit point, $\hat{f}$, until the likelihood has reduced by 0.5.

5.3 The Likelihood Ratio

It is sometimes useful to compare the likelihood for two different situations. Comparisons can be made using a likelihood ratio:

$$\lambda = 2\left(\mathbb{L}(A) - \mathbb{L}(B)\right) \tag{5.13}$$

where $\mathbb{L}(A)$ and $\mathbb{L}(B)$ are the maximum log-likelihood for two different models A and B respectively. The likelihood ratio statistic, λ, is expected to follow a χ^2 distribution.

5.4 Probability Density Functions

PDFs can be created either from binned events, or from analytical functions. Binned PDFs are easily formed from simulated events or calibration data, and can be trivially created in multiple dimensions to account for correlations between observable parameters. The Monte Carlo used to simulate SNO events is well-calibrated and should provide a good estimate of what each event distribution looks like. However, if an insufficient number of events is used to create a binned PDF, statistical fluctuations may distort the predicted distribution and lead to biases in the likelihood fit. Therefore, an appropriate division of bins is required: if the PDF bins are too wide, useful information could be excluded from the fit, but if they are too narrow, these statistical effects come into play. Multi-dimensional PDFs are much more sensitive to such statistical limitation. The discrete nature of binned PDFs also complicates procedures when fitting for systematic parameters, as described in chapter 7.

Analytical parameterisations for the observable distributions in SNO are less computationally intensive to use in the likelihood fit, but uncertainty in these approximate parameterisations can introduce additional systematic uncertainty. The analysis code described in section 5.5 was developed to use either type of PDF, or a combination of the two but binned PDFs were selected for the analysis presented in this thesis.

Each PDF must be normalised to unity over the exact data range for a given observable so that the information from each PDF has equal relevance.

5.4.1 Correlated Observables

Ideally, multi-dimensional PDFs would be used to return the probability, $F(i/j_1, j_2, j_3, ...)$ that an event belongs to signal, i, given the values of each observable, j, used in the fit. However, when a number of observables are used this can result in large, multi-dimensional distributions which can be computationally impractical to produce with sufficient statistical sampling of the distribution. If the distributions in the different observables are not correlated, the PDFs can be factorised into separate, one-dimensional PDFs as shown in equation 5.14.

$$F(i/j_1, j_2, j_3) = A(i/j_1) \times B(i/j_2) \times C(i/j_3) \tag{5.14}$$

However, if observables are correlated, this factorisation can lead to biases. This is illustrated in figure 5.1, which shows the 2-dimensional contour plot and 1-dimensional projections in x and y for two Gaussian observables. For the uncorrelated observables on the right of this figure, the projection onto the y axis gives a fair representation of the probability for any value on the x axis. This is not the case when observables are

correlated as shown on the left of the figure. For example, given the marked x-position, the projection onto the y-axis does not give the correct probability density distribution, and thus incorrectly estimates the probability at the given position.

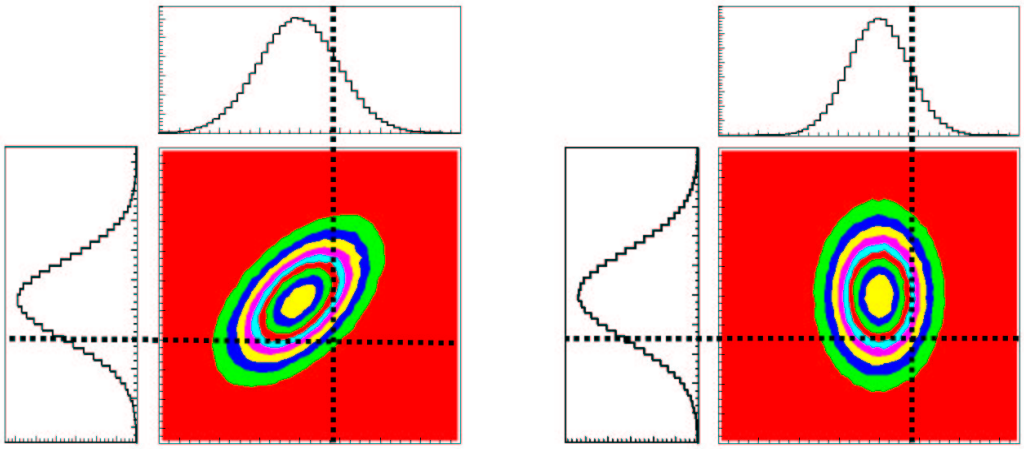


Figure 5.1: Illustration of correlated (left) and uncorrelated (right) observables with their 1-dimensional projections onto the x and y axes.

5.4.2 Conditional PDFs

For the purpose of factorisation, conditional PDFs may sometimes be used. For example, if observable j_1 is correlated with both j_2 and j_3, and observable j_4 is correlated with only j_1, the probability for an event to belong to signal i can be factorised as given in equation 5.15.

$$P(i/j_1, j_2, j_3, j_4) = P(i/j_1, j_2, j_3).P(i/j_4/j_1) \tag{5.15}$$

in which $P(i/j_1, j_2, j_3)$ is obtained from a 3 dimensional PDF, and $P(i/j_4/j_1)$ is obtained from a 2 dimensional conditional PDF. This conditional PDF is normalised separately for each bin in j_1.

5.5 The MXF Code

All signal extraction discussed in this manuscript was carried out using the MXF [84] processor: a Fortran based signal extraction code, which was developed over the course of this thesis project. This processor, which was built into the existing framework of the SNOMAN analysis code, determines the set of fit parameters (and their respective

uncertainties) which maximise the likelihood function for a given data set. The following chapters introduce different approaches to fitting the salt data, each requiring different functionality. A high level of flexibility was incorporated into the MXF code so that it can be readily applied to different fit scenarios and to different data sets. This section details some of these features and others will be introduced in subsequent chapters.

5.5.1 Cross-Checks

A number of cross-checks were built into the MXF code. A range of different minimisation algorithms, which operate on the negative likelihood value are available. The fastest algorithm is the MINUIT [83] fitter, which can be used to provide both errors on the fit parameters and a correlation matrix. However, other algorithms [85, 86] provide a valuable cross-check of the validity of the fit and uncertainty estimates.

Another useful feature is the ability to explicitly extract likelihood values at defined points in the multi-dimensional parameter space around the best fit to check that the likelihood space behaves smoothly. To scan the likelihood space in dimension X, fit parameter x is moved away from the original fit values in fixed steps and the maximum log-likelihood, $\mathbb{L}(x)$, is obtained by varying all other fit parameters at each step.

It is necessary to provide the fitter with initial values for each fit parameter. Although, in principle, the starting point should not affect the outcome of the fit, for some cases the likelihood space does not follow the smooth distribution expected and fit algorithms can be fooled by local minima. Repetition of the MXF fit for a range of different starting points guards against this.

5.5.2 PDF Weighting Factors

Nominally, each simulated event assigned to a PDF is given a weight, $W = 1$. However, W can be modified based on the observable values of an event. One particular application of this feature is the creation of PDFs for different MSW oscillation scenarios.

MSW Weighted PDFs

Some models for flavour conversion predict a distortion in the energy spectrum of electron-type neutrinos arriving at Earth, which would affect the shape of the energy PDFs. The distortion due to the MSW effect depends on the values of the mixing parameters Δm^2 and $\tan^2\theta$. If an energy PDF is used in signal extraction, an explicit assumption must be made about the model. Different models can be tested by performing the maximum likelihood fit with the relevant energy PDFs. The relative probability of these models is then obtained from a likelihood ratio as given in section 5.3.

The MXF code uses tabulated MSW suppression factors for a range of mixing parameters (Δm^2 and $\tan^2\theta$) and neutrino energies. These factors were calculated by the QPhysics processor, which is a C++ MSW analysis package developed at Queen's University that has been incorporated into SNOMAN. For a given MSW model, W is multiplied by the survival probability relating to the generated neutrino energy of each simulated event. Section 8.6 shows how this approach can be used to scan MSW parameter space.

The survival probability weighting can also be used to distort simulated data sets for verification purposes.

5.6 Application to SNO Data

5.6.1 The Salt Data Set

The salt data discussed in this thesis was collected between July 2001 and August 2003. To ensure neutrino data quality, each selected run must satisfy a number of requirements as detailed in appendix B. The selected data set consisted of 1212 runs, with a livetime of 391.43 days [87]. This livetime, and all livetimes discussed in this thesis, have been corrected for the fraction of data removed by the muon-follower cut, which is described in appendix C. These data were subjected to a number of event selection criteria, which were designed to remove background events, both instrumental and radioactive, from the data sample, as detailed in appendix C. The resulting number of candidate neutrino events in the salt data set was 4750.

5.6.2 Backgrounds Included in the Likelihood Fit

Not all background events can be removed by selection criteria and must be taken into account either during or after the signal extraction procedure. A complete discussion of the treatment of different background classes is given in 6.4. In addition to the three neutrino signals, CC, ES and NC, two classes of background were included in the likelihood fit.

The first was External Neutrons (EN events), created outside the D_2O volume by mechanisms such as (α, n) reactions in the acrylic vessel which can nonetheless enter the fiducial volume. The number of external neutron events was a parameter of the fit.

The second class of background with a characteristic observable distribution was *internal* γ events. Such γs can be created inside the D_2O by spontaneous fission of ^{238}U, or atmospheric neutrino interactions and have a characteristic distribution in energy. As the contribution of this background is small it is sufficient to fix the magnitude of this signal in the likelihood fit to the predicted value.

5.6.3 PDFs

Neutrino events generated by the SNOMAN Monte Carlo simulation package were used to create binned PDFs for the salt data analysis. The simulated detector configuration was appropriately modified for each run according to the number of operational channels and the stored calibration constants. The number of events generated for each signal corresponded to 200 times the Standard Solar Model prediction [23] over the livetime of the salt data set.* The selection criteria applied to the simulated events are also detailed in appendix C.

To create PDFs for EN events, Monte Carlo simulations for neutrons at the acrylic vessel were generated, whilst PDFs for internal γ events were created from simulations of ^{16}N events, which have a γ energy of 6.13 MeV.

Figure 5.2 shows one dimensional projections of the PDFs used for signal separation of the salt data. The observable parameters chosen for the PDFs were $\cos\theta_{\odot}$, T_{eff}, R^3 and β_{14}. R^3, is defined as:

$$R^3 = \left(\frac{R_{\rm V}}{600\,{\rm cm}}\right)^3 \tag{5.16}$$

where $R_{\rm V}$ was the reconstructed radius of the event vertex, and 600 cm is the radius of the AV. The parameter, β_{14}, measures the isotropy of an event and is described further in appendix D. The $\cos\theta_{\odot}$ distribution is required for separation of the ES distribution; R^3 information gives little separation between the three signals from neutrino events but was required for separation of EN events.

The effective electron kinetic energy distribution, T_{eff}, can separate signals, although the shape of this distribution may change for the CC and ES signals under different ν mixing scenarios. Therefore, using this distribution in the fit makes explicit assumptions about the oscillation model and the fit performed is "energy-constrained". An alternative is to perform an "energy-unconstrained" fit in which the number of CC and ES events in each energy bin is treated as a separate fit parameter as discussed in section 6.6. The distributions shown in figure 5.2 correspond to an undistorted ^{8}B spectrum [29].

5.6.4 Correlations Between Observable Parameters

The isotropy distribution is energy-dependent as higher energy electrons have straighter tracks, which produce better defined Čerenkov cones. In addition, events resulting from neutron captures that produce more γs, yield less Čerenkov light because more energy is lost in Compton scattering. Therefore, the correlation is stronger for NC and EN signals

*For ^{8}B neutrinos the scale factor was 200, for *hep* neutrinos the scale factor used was 2000.

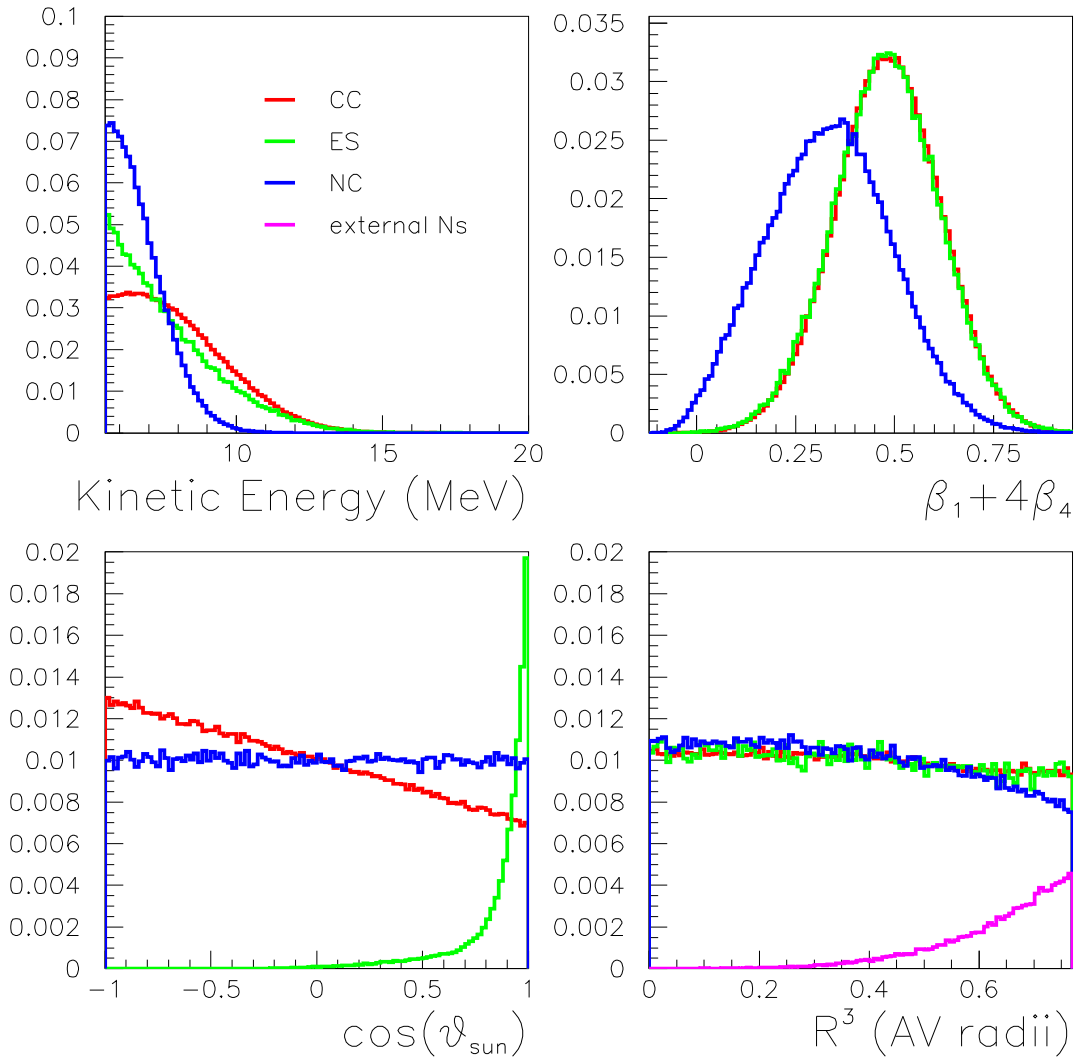


Figure 5.2: The observable distributions used for signal separation in the salt phase of SNO. These distributions were obtained from high statistics Monte Carlo simulations and are shown with an arbitrary normalisation. As EN events only differ appreciably from the NC signal in R^3, distributions in the other observables have been omitted for this signal.

than for the CC and ES signals. To account for this correlation, the observables T_{eff} and β_{14} were combined in the same multi-dimensional PDF.

The distribution of $\cos\theta_\odot$ also exhibits some energy dependence as higher energy electrons produce more light and also scatter less, so their direction can be reconstructed more accurately. For ES events, electrons travelling in the forward direction ($\cos\theta_\odot = 1$) carry away the most energy from the incoming neutrino and therefore, the width of the $\cos\theta_\odot$ peak is narrower for events with higher energy.

There is also some indication of a correlation between β_{14} and radial position [57]. This is not entirely surprising as the isotropy of an event is calculated using the fitted event vertex position. Similarly, calculation of $\cos\theta_\odot$ requires the fitted event direction.

Correlations between these reconstruction parameters have been omitted from previous analyses [28, 30, 41] as the effects are small. However, these effects are still noticeable and statistically significant biases were observed in tests of the MXF code when only the correlation of T_{eff} and β_{14} was taken into account. Section 5.7.1 details an investigation of the biases introduced into signal extraction by these correlations.

5.6.5 Bin Sizes

Table 5.1 details the selected bin divisions used for each of the four observables. These selections were based on the detector resolution in each variable, and the number of events expected across the parameter range. For example, the SNO detector energy resolution, given by [88]:

$$\sigma_{\mathrm{E}} = -0.131 + 0.383\sqrt{T_{\mathrm{gen}}} + 0.03731 T_{\mathrm{gen}}, \tag{5.17}$$

where T_{gen} is the generated kinetic energy, is 1–2 MeV in the selected energy range. Based on this knowledge, 0.5 MeV wide bins were selected for the T_{eff} PDF. However, above 13.5 MeV, very few neutrino events are expected. Therefore, to prevent statistical fluctuations influencing the fit in this high energy region, a single 6.5 MeV wide bin was used. An investigation into the effect of changing bin sizes for all the selected observables is described in appendix E.

5.7 Verification of the Signal Extraction Code

In order to test the signal extraction technique against possible biases, tests were carried out on 100 artificial data sets, created from half of the simulated events. The remaining simulated events were reserved to create the PDFs. As simulated events are generated on a run-by-run basis to model detector conditions as closely as possible, it was necessary to sample evenly over time when creating the artificial data sets. Initially, these data sets were created with a fixed number of each type of event and more parameters were fitted

Observable	Minimum	Maximum	Detector Resolution	Bin sizes
$\cos\theta_\odot$	-1.0	1.0	$\approx 10^\circ$ in $\theta_\odot$	100 × 0.02
R^3	0.0	$\left(\frac{550}{600}\right)^3$	$\approx$ 20 cm in R	10× 0.077
T_eff	5.5 MeV	20.0 MeV	1-2 MeV	16 × 0.5, 1 × 6.5 MeV
β_{14}	-0.12	0.95	-	50 × 0.0214

Table 5.1: Parameter ranges and binning used for each observable in the salt PDFs. The approximate values for the detector resolution should be viewed only as a rough guide.

Signal	Mean Generated Events
CC	2027 ± 5.02
ES	309 ± 1.73
NC	2189 ± 4.56
EN	149 ± 1.30

Table 5.2: Mean numbers of events generated for the 100 artificial data sets used for verification of the signal extraction procedure.

within 1σ of their true values than expected. This is because the extended likelihood approach assumes Poissonian fluctuations in the number of events in the data set. To approximate a Poissonian distribution for the simulated data sets, data selection criteria were not applied until after events had been assigned to a data set, to randomly adjust the number of "good" events. The mean generated number of each signal type is given in table 5.2.

For the energy-constrained fit considered here, there were four variable parameters in the likelihood function and each fit was completed in a few minutes, but as the number of parameters is increased the fitting procedure becomes more time consuming. For this reason, tools were established in order to run the fits on the simulated data sets simultaneously on a large CPU farm. A series of Perl scripts were written to initiate each fit and combine the results for each scenario tested.

The total fitted number of CC, ES, NC and EN events in each case was compared to the expected number of events. For each data set, the bias for fit parameter X was defined by:

$$Bias(X) = \frac{F(X) - E(X)}{\sigma(X)} \tag{5.18}$$

where $\sigma(X)$ is the statistical uncertainty on the fitted parameter, $F(X)$, from the maximum likelihood fit and $E(X)$ is the expected number of events of type X. The fractional

deviation for fit parameter X, is similarly defined by:

$$\Delta(X) = \frac{F(X) - E(X)}{E(X)} \tag{5.19}$$

The standard likelihood function used for the verification tests described in this section was:

$$\mathbb{L} = \sum_{d=1}^{N_{\rm data}} log \left(\sum_{i=\rm CC,ES,NC,EN} n_i P(i/T_{eff}, \beta_{14}, R^3, \cos\theta_\odot) \right) - \sum_{i=\rm CC,ES,NC,EN} n_i \tag{5.20}$$

where $P(i/T_{eff}, \beta_{14}, R^3, \cos\theta_\odot)$, is the combined probability for signal type i, which can be factorised in a number of ways as described below. Since the data and PDFs were simulated with the same energy distributions, the energy shape was constrained in the fit through the use of T_{eff} information for the purposes of this study.

5.7.1 Sensitivity to Correlations of Observable Parameters

Factorisation of the probability allows the used of PDFs with lower dimensionality, which are less susceptible to statistical fluctuations. Different factorisations of the probability function were tested against bias. In all these tests, PDFs were created using the binning specified in table 5.1. The first factorisation for probability was:

$$P(i/T_{eff}, \beta_{14}, R^3, \cos\theta_\odot) = F(i/T_{eff}, \beta_{14}).F(i/R^3).F(i/\cos\theta_\odot) \tag{5.21}$$

Here only the strongest correlation between energy and isotropy distributions was taken into account with a two dimensional PDF. Both the R^3 and $\cos\theta_\odot$ probabilities were obtained from one dimensional PDFs. The results of fits to the artificial data sets using this function are given in table 5.3. In this scenario, the bias seen for all the fitted signals differs from zero with at least 3σ significance. Whilst this contributes less than 1% fractional uncertainty to the CC and NC fitted values, the effect is more significant in the ES signal and substantial for the EN events.

As the EN signal only differs significantly from the NC signal in the radial distribution, the biases seen in the above test were attributed, at least partly, to correlations with R^3 which are not taken into account with the one dimensional R^3 PDF. To test this hypothesis, a second test was carried out with probabilities factorised as given in equation 5.22, using a three dimensional PDF:

$$P(i/T_{eff}, \beta_{14}, R^3, \cos\theta_\odot) = F(i/T_{eff}, \beta_{14}, R^3).F(i/\cos\theta_\odot) \tag{5.22}$$

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2010.6 ± 6.9	-0.23 ± 0.06	-0.008 ± 0.002	68.584 ± 0.096
ES	317.1 ± 2.9	0.30 ± 0.07	0.030 ± 0.007	29.354 ± 0.067
NC	2168.4 ± 7.9	-0.26 ± 0.08	-0.009 ± 0.003	78.594 ± 0.092
EN	173.7 ± 4.5	0.56 ± 0.09	0.171 ± 0.026	45.217 ± 0.073

Table 5.3: Results for fits to 100 artificial data sets using probabilities factorised as given in 5.21, and the likelihood function given in 5.20.

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2030.4 ± 6.8	0.06 ± 0.06	0.002 ± 0.002	68.467 ± 0.094
ES	310.5 ± 2.9	0.07 ± 0.07	0.008 ± 0.007	29.133 ± 0.067
NC	2168.0 ± 7.7	-0.27 ± 0.08	-0.009 ± 0.003	77.012 ± 0.096
EN	159.8 ± 4.4	0.26 ± 0.09	0.079 ± 0.027	44.150 ± 0.082

Table 5.4: Results for fits to 100 artificial data sets using probabilities factorised as given in 5.22, and the likelihood function given in 5.20.

Correlations with $\cos\theta_\odot$ were not accommodated in this scenario as a one-dimensional PDF was still used for this observable. The results of fits to 100 artificial data sets using this function are given in table 5.4. The biases on the CC and ES signals are dramatically reduced by the use of the three dimensional PDF, and the bias in the EN signal has less than 3σ significance. The events incorrectly assigned to the EN signal appear to be drawn mostly from the NC signal, causing a 0.9% bias. This bias is small compared to other possible systematic effects on the fitted NC signal, which are discussed in chapter 6.

To study correlations with the $\cos\theta_\odot$ distribution, two more factorisations of the likelihood function were tested. Firstly correlations of $\cos\theta_\odot$ with reconstructed vertex position were taken into account by the factorisation given in 5.23.

$$P(i/T_{eff}, \beta_{14}, R^3, \cos\theta_\odot) = F(i/T_{eff}, \beta_{14}, R^3).F(i/\cos\theta_\odot/R^3) \tag{5.23}$$

Whilst the same three-dimensional PDFs were used to provide $F(i/T_{eff}, \beta_{14}, R^3)$, the probability that an event with a given $\cos\theta_\odot$ value is drawn from distribution i was obtained from a PDF that is conditional on that event's R^3 value. The results of fits to the artificial data sets using this function are given in table 5.5. The observed biases are reduced slightly by accounting for this additional correlation.

It should be noted that the energy-unconstrained likelihood fit carried out in chap-

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2027.9 ± 6.8	0.03 ± 0.06	0.001 ± 0.002	68.468 ± 0.094
ES	310.3 ± 2.9	0.07 ± 0.07	0.008 ± 0.007	29.141 ± 0.067
NC	2172.7 ± 7.7	-0.21 ± 0.08	-0.007 ± 0.003	77.221 ± 0.095
EN	157.9 ± 4.4	0.22 ± 0.09	0.066 ± 0.026	44.376 ± 0.080

Table 5.5: Results for fits to 100 artificial data sets using probabilities factorised as given in 5.23, and the likelihood function given in 5.20.

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2038.2 ± 6.8	0.18 ± 0.06	0.006 ± 0.002	68.188 ± 0.094
ES	307.1 ± 2.9	-0.04 ± 0.08	-0.003 ± 0.007	29.058 ± 0.068
NC	2166.8 ± 7.7	-0.28 ± 0.08	-0.010 ± 0.003	77.000 ± 0.095
EN	155.7 ± 4.4	0.17 ± 0.09	0.052 ± 0.026	43.986 ± 0.079

Table 5.6: Results for fits to 100 artificial data sets using probabilities factorised as given in 5.24, and the likelihood function given in 5.20.

ter 6 naturally accounts for correlations between $\cos\theta_\odot$ and T_{eff}, as separate PDFs for $\cos\theta_\odot$ are constructed for each T_{eff} bin. However, the energy-constrained fits tested in this chapter were still susceptible to bias due to this correlation. The effects were studied by factorising the probability distributions as given in 5.24.

$$P(i/T_{eff}, \beta_{14}, R^3, \cos\theta_\odot) = F(i/T_{eff}, \beta_{14}, R^3).F(i/\cos\theta_\odot/T_{eff}) \tag{5.24}$$

This scenario was similar to the previous one, but with the $\cos\theta_\odot$ PDF conditional on the event energy, T_{eff}, rather than R^3. Results from this set of fits are given in table 5.6. As expected, this approach reduces the bias in the fitted ES signal to an insignificant level. However, the biases in the CC and NC signals are larger than in the previous scenario, possibly due to the correlation between $\cos\theta_\odot$ and R^3 that was not treated here.

5.8 Summary

This chapter introduced extended maximum likelihood, a powerful statistical technique that is ideally suited to separation of the different types of event seen in SNO. The likelihood function can be parameterised to include information obtained from a number of characteristic observable distributions and additional constraints on the parameters of the fit. The method has been shown to be extremely sensitive to low-level correlations between

the observable distributions. The MXF code has been developed with a high degree of flexibility and such correlations can be accounted for through multi-dimensional PDFs.

The tools introduced in this chapter have been extensively tested on simulated data sets and perform well for energy-constrained fits. In the next chapter, the MXF code will be used to perform an energy-unconstrained fit to the salt data set in order to measure the charged-current energy spectrum for solar neutrinos.

Chapter 6

The Charged-Current Energy Spectrum

MSW mixing can lead to a distortion of the electron neutrino energy spectrum which, in turn, affects the shape of the T_{eff} distribution for the CC and ES signals (see figure 1.5). Thus, a direct measurement of the CC energy spectrum can be used to test MSW oscillation models directly. To interpret such a measurement, a full understanding of the uncertainties contributed to the measured spectrum by systematic effects is required. In this chapter, these systematic effects are discussed and a measurement of the CC energy spectrum in T_{eff} is presented.

6.1 Systematic Effects

Systematic differences between data and the simulated events used to create PDFs will affect the results of a likelihood fit. Whilst every effort is made to model the detector response as precisely as possible, small differences in the predicted event characteristics cannot be ruled out. However, possible systematic effects can be characterised through detailed comparison of calibration data and simulated events.

The results of these studies can be used to make corrections to the predicted distributions of observables and to place bounds on the systematic uncertainty associated with them. A number of the systematic effects could be energy-dependent and, therefore, must be treated differentially in the context of a spectral analysis.

The systematic effects considered for the purpose of this thesis can be divided into two types: those that modify the predicted values of observable parameters (namely T_{eff}, β_{14}, $\cos\theta_{\odot}$ and R^3 as defined in section 5.6.3) on an event-by-event basis and those which are not associated with a specific observable, but still affect the differential amplitude of probability distributions. The former were treated by adjusting the observable parameter values of simulated events, whilst the latter were taken into account with PDF weighting

factors (as introduced in section 5.5.2). These two types of systematic are discussed in sections 6.2 and 6.3, respectively.

6.2 Observable Parameter Systematics

For the spectral analysis, a range of calibration data was used to characterise systematic effects at different energies. As there is only limited calibration data sampling the range $9 \lesssim T_{eff} \lesssim 20\,\mathrm{MeV}$, the detector model is less well constrained in this region, which is reflected by larger systematic uncertainties. This section details the evaluation and treatment of the systematic effects on observable parameters that contribute to uncertainties in the CC spectrum.

6.2.1 Uncertainty in T_{eff}

A number of factors contribute to uncertainty in predictions of the observable effective kinetic energy parameter, T_{eff} (defined in section 2.4.2). For example, systematic differences in the modelled energy response of the detector could arise due to a drift in the energy response with time, poor modelling of PMT gain and threshold effects or incorrect treatment of non-operational PMTs.

These effects were studied in detail using the ^{16}N calibration source. Possible biases in the energy response with detector radial position were also studied using ^{16}N data obtained at a range of positions in the D_2O volume. As ^{16}N data was collected at a higher event rate ($\approx 200\,\mathrm{Hz}$) than neutrino data ($\approx 20\,\mathrm{Hz}$), there is an additional uncertainty associated with the detector response because the calibration constants for each channel are dependent on the rate at which it is read out [89]. There is also an additional uncertainty in the evaluated systematic effects due to uncertainty in modelling of the ^{16}N source.

A difference in the energy-mapping between data and simulated events was parameterised by a scale factor α that maps the T_{eff} value for simulated events to

$$T'_{eff} = T_{eff}(1 + \alpha) \tag{6.1}$$

in order to replicate the data distribution in T_{eff}. Comparisons of the ^{16}N data with simulated events resulted in a value of α consistent with zero with an uncertainty of $\sigma_{\mathrm{Tsc}} = \pm\mathbf{1.13\%}$ [88]. Resolution was defined as the spread of the T_{eff} distribution for events of fixed energy. Similar studies of ^{16}N indicated that the simulation under-estimated the resolution in T_{eff} by $\delta_{\mathrm{Tres}} = \mathbf{1.8\%}$. In order to correct the simulated T_{eff} distribution, the following expression was used to modify the value of T_{eff} for each simulated event

$$T'_{eff} = T_{eff}(1 + g(\delta_{\mathrm{Tres}})) \tag{6.2}$$

where $g(\delta_{\rm Tres})$ is a value randomly selected from a Gaussian distribution with width $\delta_{\rm Tres}$ and zero mean. The 1σ uncertainty on this systematic was parameterised by

$$\sigma_{\rm Tres} = 1.6 + 0.478(T_{eff} - 5.5\,{\rm MeV})\% \quad \text{for} \quad T_{eff} > 5.5\,{\rm MeV}. \tag{6.3}$$

In addition to these systematic effects, the difference between the observed and simulated T_{eff} values for ^{16}N were observed to change with the radial position of the ^{16}N source. This additional radial-dependent systematic in the simulated T_{eff} distribution was corrected for with a further modification to the T_{eff} value of each simulated event:

$$T'_{eff} = T_{eff}\left(1.0 - 0.01\left(\frac{R}{550\,{\rm cm}}\right)\right) \tag{6.4}$$

where R is the reconstructed event vertex. The uncertainty on this systematic was negligible.

These three systematic estimates were based solely on comparisons of the ^{16}N data and simulations. For the spectrum analysis, an additional *non-linearity* uncertainty was assigned to account for an energy-dependent change in the difference between simulated and observed T_{eff}:

$$\sigma_{\rm Tnl} = 0.09(T_{eff} - 5.05\,{\rm MeV})\% \tag{6.5}$$

The value of 5.05 MeV corresponds to the kinetic energy of an electron that produces the same mean T_{eff} value as ^{16}N events. The study carried out to determine the magnitude of this systematic uncertainty is described in detail in appendix G.

6.2.2 Uncertainty in β_{14}

The systematic effects associated with the isotropy parameter, β_{14}, (defined in section D.1) were determined from the fractional differences measured between β_{14} distributions for calibration data and simulated events. These systematics were parameterised in terms of changes in the mean and width of the β_{14} distribution. For reference, the mean and width of the β_{14} distribution for each signal is given in table 6.1. These values were obtained from Gaussian fits to the β_{14} distribution of simulated events. Whilst the CC and ES distributions fit well to a Gaussian, the NC and EN distributions are slightly skewed, and the Gaussian approximation is not perfect.

As the inherent event isotropy is broader for events observed due to neutron capture interactions (*i.e.* NC and EN events) than for events observed by Čerenkov light emitted directly by electrons produced in the interaction (*i.e.* CC and ES events) systematics are likely to have a different effect on the β_{14} distributions. It is also conceivable that the "neutron-type" events are sensitive to different systematic effects to the "electron-type" events. For example, uncertainties in the modelling of the γ cascade produced by neutron

Signal	β_{14} mean	β_{14} width
CC	0.483	0.133
ES	0.479	0.135
NC	0.335	0.157
EN	0.346	0.153

Table 6.1: Mean and width of the β_{14} distribution for each type of neutrino signal integrated over the energy range used for analysis.

capture on chlorine, or in the angular distribution for Compton scattering, would only affect the former type of event. Therefore, the systematic differences in β_{14} between data and simulated events were characterised separately for electron-type and neutron-type events.

Incorrect simulation of the energy-dependence of the isotropy distribution could also have a significant effect on the results of the likelihood fit. Characterisation of energy-dependent systematic effects is especially important in the energy-unconstrained fit for the CC spectrum. In the absence of T_{eff} information, the β_{14} distribution dominates the separation of CC and NC-type events.

Studies of ^{16}N and ^{8}Li data [57, 90, 91, 92] to characterise systematic effects on the β_{14} distribution of electron-type events indicated a difference between the observed and predicted mean value of β_{14} which could be parameterised by:

$$\delta_{\rm be} = a_{\rm e} + b_{\rm e}(T_{eff} - 5.05\,{\rm MeV}) \tag{6.6}$$

where the parameters $a_{\rm e}$ and $b_{\rm e}$ take the values given in table 6.2. $a_{\rm e}$ was derived from differences between data and simulation observed for the ^{16}N source. Since calibration data was only collected at a set of fixed points in detector space the main uncertainty on this parameter was due to possible variations in measured β_{14} with position. This uncertainty was determined from the RMS of the mean β_{14} values obtained from all source positions. $b_{\rm e}$ was derived from linear fits to data-simulation differences for ^{8}Li. These fits were constrained by values obtained for ^{16}N data for the same detector locations.

To correct for this systematic effect, the β_{14} value for each simulated CC or ES event was modified by

$$\beta'_{14} = \beta_{14} + \bar{\beta}_{14}\delta_{\rm be} \tag{6.7}$$

where $\bar{\beta}_{14}$ is the mean value of the β_{14} distribution for the simulated event-type. The uncertainty on $\delta_{\rm be}$ was taken to be:

$$\sigma_{\rm be} = \left(da_{\rm e}^2 + (db_{\rm e}(T_{eff} - 5.05\,{\rm MeV}))^2\right)^{1/2} \tag{6.8}$$

Parameter	Value	Uncertainty
$a_{\rm e}$	0.00186	$\pm$ 0.00852
$b_{\rm e}$	0.00085	$\pm$ 0.00077
$c_{\rm e}$	0.00929	$\pm$ 0.00657
$a_{\rm n}$	-0.003	$\pm$ 0.00852
$b_{\rm n}$	0.0037	$\pm$ 0.00047
$c_{\rm n}$	0.00555	$\pm$ 0.00657

Table 6.2: Parameters describing systematic effects on β_{14} parameters for both electron-type and neutron-type events.

where $da_{\rm e}$ and $db_{\rm e}$ are the uncertainties on the parameters $a_{\rm e}$ and $b_{\rm e}$ also given in table 6.2.

The systematic effect on the width of the β_{14} distribution for electron-type events was determined from a similar study of ^{16}N. To correct for differences between data and simulation, the value of β_{14} for each simulated CC or ES event was further modified by:

$$\beta'_{14} = \beta_{14} + (\beta_{14} - \bar{\beta}_{14})c_{\rm e} \tag{6.9}$$

where the value of $c_{\rm e}$, and the associated uncertainty, $dc_{\rm e}$ are also given in table 6.2.

The systematic effect on the mean β_{14} value for neutrons was obtained from a linear fit to the difference between ^{252}Cf data and simulations:

$$\delta_{\rm bn} = a_{\rm n} + b_{\rm n}(T_{eff} - 5.05\,{\rm MeV}) \tag{6.10}$$

Again, the parameters derived from this fit are given in table 6.2. As ^{252}Cf calibration data samples the detector even more sparsely than ^{16}N data, the uncertainty on $a_{\rm n}$ was taken to be the uncertainty due to spatial variations obtained from the ^{16}N data. Simulated NC and EN events were corrected by $\delta_{\rm bn}$ as described for electron-type events in equation 6.7. Similarly, systematic differences between the width of data and simulated β_{14} distributions were corrected for by modifying the β_{14} value of NC and EN events by $c_{\rm n}$ (also given in table 6.2) as described in equation 6.9. This systematic was also assigned a spatial sampling uncertainty obtained from ^{16}N. The uncertainties on these two corrections, $\sigma_{\rm bn}$ and $dc_{\rm n}$, were parameterised in the same way as the electron-type β_{14} systematic uncertainties.

6.2.3 Uncertainty in $\cos\theta_{\odot}$

The direction of each observed event is reconstructed based on the timing and angular distribution of PMT hits with respect to the fitted vertex position. Systematic effects

on the reconstructed direction could result from a systematic change in the fitted vertex position or an error in the angular distribution assumed for the PMT hits, and would affect the parameter $\cos\theta_\odot$, the cosine of the angle between reconstructed event direction and the direction to the Sun.

Details of studies of the angular response of the SNO detector can be found in Ref. [93]. The uncertainty in the parameters characterising the angular response was evaluated through comparison of the angular distributions for ^{16}N data and simulated events. A change in the $\cos\theta_\odot$ distribution, comparable to the magnitude of this uncertainty, was replicated by adjusting the direction of each event by an angle $\theta_{\rm R}$. This energy-dependent angle was selected from a Gaussian distribution in $\cos\theta_{\rm R}$ with width, W:

$$W(\cos\theta_{\rm R}) = e^{-2.935-0.145\times T_{eff}} \tag{6.11}$$

The smeared event direction was selected randomly from a cone of angle $\theta_{\rm R}$ around the original event direction, resulting in a new value of $\cos\theta_\odot$.

6.2.4 Uncertainty in R^3

The position of each observed event is reconstructed based on the timing distribution of PMT hits, and is described by the parameter R^3, specified in terms of cubic AV radii (see equation 5.16). For CC, ES and NC signals the distribution in this parameter is nearly flat, so uncertainties in event position are only expected to affect the shape of the EN PDF. Offsets in the reconstructed x, y or z position are linear effects, which will not alter the overall fiducial volume and are therefore likely to have little affect on signal extraction. Such offsets could result from a poor knowledge of the calibration source position, or anisotropy in the source event distribution. A difference in the radial *scale* between data and simulation is a more serious effect that would change the number of events in the fiducial volume. Differences in radial scale could result from errors in the simulated refractive index, or optical features of the detector that are not precisely modelled and would be apparent through a radial dependence of the difference between reconstructed and expected vertex position.

Studies of vertex reconstruction indicated a movement of the acrylic vessel in the salt phase of the experiment, presumably due to the addition of an extra 2 tonnes of mass to the AV in the form of NaCl. Fits for the acrylic vessel position indicated that it was 5.5 cm lower than the D_2O phase of the experiment [94]. To account for this the reconstructed z coordinate for all data events was modified by:

$$z'_{\rm data} = z_{\rm data} + 5.5\,{\rm cm}, \tag{6.12}$$

Parameter	Systematic	Value
X	offset	$dx = \pm 2$cm
Y	offset	$dy = \pm 2$cm
Z	offset	$dz = \pm 6$cm
R	scale	$\delta_R = \pm 1.0\%$
R	resolution	$\omega_R = 0.015 \times 16\,\mathrm{cm}$

Table 6.3: Uncertainties in reconstructed event position obtained from [95].

and the reconstructed radius for all simulated events was modified by:

$$R'_{\mathrm{MC}} = 0.995 R_{\mathrm{MC}}. \tag{6.13}$$

Systematic differences between data and simulated events were then studied using these corrected reconstructed event positions for ^{16}N [95]. The uncertainties determined for the reconstructed x, y and z positions and radial accuracy are given in table 6.3. The resolution uncertainty also given in this table accounts for a possible increase in the width of the reconstructed radius distribution for events simulated at a given detector position. The factor ω_R gives the width of the Gaussian function, $g(\omega_R)$, used to modify the reconstructed radius (as for the T_{eff} resolution given in equation 6.2). This function assumes a nominal vertex resolution of 16cm, with a 1.5% uncertainty.

A change in the accuracy of the reconstructed event position with energy would cause an energy-dependence in the fraction of events inside the fiducial volume. As higher energy events produce more PMT hits and, therefore, more information for event reconstruction, such effects are expected. However, if the energy-dependence modelled differs to that observed, the spectrum measured is subject to an additional energy-dependent systematic, which is discussed in detail in appendix H and included as a PDF weighting factor (see section 6.3).

6.2.5 Corrections to Observable Parameters

Some of the studies described above produced evidence for a systematic difference between data and simulation. For the analysis presented in this chapter, corrections were made to the observable distributions to account for these systematic effects. The corrections applied to the observable parameters of simulated events are summarised in table 6.4.

Observable Parameter	Corrections
T_{eff}	resolution (6.2) and radial bias (6.4)
(R^3)	AV shift (6.13)
β_{14}	mean (6.7) and width (6.9)

Table 6.4: The corrections applied to observable parameters for simulated events to account for systematic differences between data and simulations. Note that the correction for R^3 is applied to R and the parameters for the β_{14} corrections are different for electron-type and neutron-type events. The numbers in parentheses indicate the equations that describe each correction.

6.2.6 Uncertainties in Observable Parameters

All the systematic effects measured were subject to uncertainties, which in turn contribute to an uncertainty on the parameters of a likelihood fit. The effect of systematic uncertainties on observable parameters was estimated by repeating the likelihood fit multiple times with PDFs created from simulated events with their observable values individually shifted to the 1σ systematic limits. The 1σ limits for all observable parameter systematics are given in table 6.5 along with the modification applied to simulated events. For resolution parameters $g(X)$ is a value sampled randomly from a Gaussian distribution of width X and zero mean. (See section 6.2.3 for the specific application to event direction that modifies the $\cos\theta_{\odot}$ distribution). Note that only positive resolutions are sampled, as it is not possible to make a distribution narrower in this manner.

Each systematic uncertainty was estimated separately and different fits were performed for positive and negative uncertainties on each parameter. For systematically corrected observables, the uncertainties were considered with respect to the corrected value.

6.3 PDF Weighting Systematics

Some systematic differences between data and simulations were taken into account using the PDF weighting factors that were introduced in section 5.5.2. This accounts for a change in PDF shape and also a change in the PDF normalisation. The energy-dependent effects detailed in this section were treated in this way.

Radiative corrections

Bremsstrahlung radiation effects which were omitted from the Monte Carlo simulation, were corrected by a factor ω_j^{RC}. For the CC and ES signals, this correction was energy-dependent as given by equations I.1 and I.2 in appendix I.

Observable	Systematic	Modification	Value(s)
T_{eff}	Scale	$T'_{eff} = T_{eff}(1 \pm \sigma_{\mathrm{Tsc}})$	1.13%
T_{eff}	Resolution	$T'_{eff} = T_{eff}(1 + g(\delta_{\mathrm{Tres}}))$	equation 6.3
T_{eff}	Non-linearity	$T'_{eff} = T_{eff}(1 \pm \sigma_{\mathrm{Tnl}})$	equation 6.5
β_{14}	Mean	$\beta'_{14} = \beta_{14} \pm \bar{\beta}_{14}\sigma_{bx}$	equation 6.8
β_{14}	Width	$\beta'_{14} = \beta_{14} \pm (\beta_{14} - \bar{\beta}_{14})dc_x$	table 6.2
$(\cos\theta_{\odot})$	Resolution	$W(\cos\theta_{\mathrm{R}})$	equation 6.11
(R^3)	X Offset	$x' = x \pm dx$	table 6.3
(R^3)	Y Offset	$y' = y \pm dy$	table 6.3
(R^3)	Z Offset	$z' = z \pm dz$	table 6.3
(R^3)	R Scale	$R' = R(1 \pm \delta R)$	table 6.3
(R^3)	R Resolution	$R' = R(1 + g(\omega_R))$	table 6.3

Table 6.5: Systematic uncertainties applied to simulated observable parameters. Parenthesis are used when the systematic is not applied directly to the observable. For the β_{14} systematics x=e for electron-type events, and x=n for neutron-type events. See section 6.2.3 for the parameter modification required for $\cos\theta_{\odot}$.

Aborted simulation events

The factor, MC_{corr}, was used to correct for aborted simulated events and is discussed in appendix F.

Energy-dependent fiducial volume uncertainty

A factor V_{corr} (see equation H.4) was used to correct for an energy-dependent difference between the number of data and simulated events that reconstruct inside the fiducial volume (discussed in appendix H).

Sacrifice

S_{ac} is a correction for the number of true neutrino events that fail the event selection criteria applied to the data set (discussed in appendix I).

All these corrections were applied in the MXF signal extraction code with a combined weighting factor:

$$W(T_{eff}) = \frac{\omega_j^{\mathrm{RC}} \left(1 - S_{ac_j}\right) V_{\mathrm{corr}}}{MC_{\mathrm{corr}}} \tag{6.14}$$

Note that all these factors are dependent on event energy, T_{eff};* the factors $\omega^{\rm RC}$ and S_{ac} are also dependent on the type of interaction, j.

The effects of uncertainty in the S_{ac} and $V_{\rm corr}$ factors were evaluated through additional likelihood fits carried out with these factors set to their $\pm 1\sigma$ limits. The effect of small uncertainty in the $MC_{\rm corr}$ factor was assumed to be negligible in the analysis presented here and uncertainties in the radiative corrections are included separately (see section 6.7.1).

6.4 Background Events

The presence of background events (see section 2.4.6 for a discussion of different background types) in the salt data set will also have a systematic effect on the number of events of each signal type. The uncertainty in backgrounds that are included as parameters of the fit (*i.e.* EN events), is naturally incorporated in the statistical uncertainties on the fitted parameters. Neutron events caused by backgrounds internal to the D_2O (INB events) are indistinguishable from NC events and, thus, are subtracted from the fitted number of NC events. The number of INB events in the salt data sample was calculated to be [49]

$$N_{\rm INB} = 138^{+30.6}_{-31.7}\,\text{events}. \tag{6.15}$$

The number of events due to the internal γ background, $N_{{\rm in}\gamma}$, was calculated to be [65]

$$N_{{\rm in}\gamma} = 8.5 \pm 2.2 \text{ events}$$

and was fixed in the likelihood fit. To determine the effect of the uncertainty in this calculation, the likelihood fit was repeated with the number of internal γ events fixed to $N_{{\rm in}\gamma} \pm 1\sigma$.

Upper limits were determined for the contribution of all other backgrounds to the salt data set. These limits, which are summarised in table 6.8, were treated as systematic uncertainties on the fitted numbers of signal events.

The limit on the number of internal Čerenkov events was calculated using:

$$N_{\rm ic}(x) = N_{\rm np}(x).t'.\epsilon_{\rm n}.R_{\rm cer}(x) \tag{6.16}$$

where, x is the isotope which decays, $N_{\rm np}$ is the production rate of neutrons from that isotope, $t' = 391.43$ days is the livetime of the salt data set [87], $\epsilon_{\rm n} = 0.398$ is the neutron detection efficiency inside all data cuts [96], and $R_{\rm cer}$ is the evaluated ratio of Čerenkov events to observed neutron events due to that particular isotope. The three isotopes which

*With the exception of $\omega^{\rm ES}$ which was defined in terms of $E_{\rm gen}$ and $MC_{\rm corr}$ which was applied to $(T_{eff}+0.511\,\text{MeV})$.

Isotope	$N_{\rm np}$ (events/day)	$R_{\rm cer}$	$N_{\rm ic}$ (events)
^{208}Tl	0.293±0.182	$0.011^{+0.005}_{-0.002}$	0.53
^{214}Bi	$0.277^{+0.046}_{-0.070}$	$0.053^{+0.011}_{-0.013}$	2.29
^{24}Na	0.064±0.016	$0.090^{+0.024}_{-0.018}$	0.90

Table 6.6: Production rate of neutrons and the ratio of Čerenkov type-events, $R_{\rm cer}$, for different radioactive isotopes. These values were used to calculate the number of internal Čerenkov backgrounds, $N_{\rm in}$ from each isotope, given in the final column.

Source	$N_{\rm ec}^X$ (events)
PMT $\beta\gamma$s	≤11.0
H_2O $\beta\gamma$s	≤3.0
AV $\beta\gamma$s	≤7.0

Table 6.7: Upper limits on the number of external Čerenkov background events from different sources in the salt data set.

can contribute to backgrounds in this way are ^{208}Tl, ^{214}Bi and ^{24}Na. The values of $N_{\rm np}$ obtained from [49] and $R_{\rm cer}$ obtained from [97, 98][†] are given in table 6.6, along with the calculated values of $N_{\rm ic}$ for each isotope. The total number of internal Čerenkov events was calculated to be

$$N_{\rm ic} = 3.55 \pm 0.79^{+0.61}_{-0.30} \text{ events.}$$

Without full knowledge of the observable distributions for these background events it was not possible to determine whether these events would likely contribute to the extracted value of CC or NC type events, although their contribution to the fitted ES flux was assumed to be negligible due to the characteristic distribution in $\cos\theta_\odot$ that separates this signal. Given that the overall number is small, the 1σ upper limit of 4.6 events was conservatively taken as a systematic uncertainty on both the number of events in the lowest CC energy bin, and the fitted number of NC events.

The limits on external Čerenkov backgrounds are given in table 6.7 and were obtained from Ref. [99]. Like the internal Čerenkov backgrounds, these events were assumed to contribute only to the lowest CC energy bin (5.5–6.0 MeV)[‡] and the fitted NC flux.

[†]For Tl and Na a Monte Carlo method was used whilst $R_{\rm cer}$ for Bi events was obtained from analysis of Radon spike data.

[‡]This assumption may not be entirely valid as the fit for PMT $\beta\gamma$ events was also sensitive to higher energy γ events created by captures in the PSUP. However, this is a second order correction which was assumed to be negligible for this analysis.

Background type	Systematic Effect (events)	Contribution
EN events	Parameter of the Likelihood Fit	
Internal Neutron events	$138^{+30.6}_{-31.7}$	Subtract from F(NC)
Internal γ events	8.5±2.2	Fixed in the Likelihood Fit
Internal Čerenkov events	≤ 4.6	F_1(CC) and F(NC)
External Čerenkov events	≤ 18.5	F_1(CC) and F(NC)
AV events	≤ 6.55	F(NC)
	≤ 3.28	F_1(CC)
	≤ 0.24	F_j(CC) $\quad j > 1$
Instrumental backgrounds	≤ 3.0	F(NC)
	≤ 0.21	F_j(CC)

Table 6.8: This table summarises the contributions of different types of background events to the salt data set. The final column indicates which fit parameter is affected by each systematic. F_j(CC) are the fitted number of events in each 0.5 MeV CC T_{eff} bin.

The combined uncertainty due to external Čerenkov events was obtained by treating the individual contributions as 1σ upper limits on a Poissonian distribution [99]:

$$N_{\rm ec} \leq 18.5 \text{ events.}$$

The number of AV events [100] was found to be

$$N_{\rm AV} \leq 6.55 \text{ events.}$$

The energy distribution of these events is not well known, though the contribution is expected to be larger at low energies. For this reason, the systematic uncertainty assigned to the first CC T_{eff} bin due to these events was $\frac{N_{\rm AV}}{2}$, whilst the uncertainty for the other 0.5 MeV bins was $\frac{N_{\rm AV}}{2\times 13.5}$. $N_{\rm AV}$ also contributed to the systematic uncertainty on the fitted number of NC events.

$$N_{\rm inst} \leq 3.0 \text{ events}$$

is the upper limit on the number of instrumental background events in the salt data set [100]. These events were assumed to have a flat distribution in energy and their systematic uncertainty contribution for each 0.5 MeV CC bin was determined to be $\frac{N_{\rm inst}}{14.5}$.

6.5 Systematic Uncertainty on Fitted Parameters

Systematic shifts in the PDFs can affect the measured spectrum in two ways. Firstly, systematic changes in the shape of the PDFs influence the number of events of each

class that maximises the likelihood function. This effect was taken into account when the likelihood fit was repeated with the systematically shifted PDFs. The uncertainty, $\delta_F^{\psi}(X)$, on a given fit parameter, $F(X)$, due to systematic shift, ψ, is taken as the difference between the fit parameter obtained with the shifted PDF, $F^{\psi}(X)$, and that obtained with the nominal fit $F^n(X)$:

$$\delta_F^{\psi}(X) = F^{\psi}(X) - F^n(X). \tag{6.17}$$

Uncertainties due to the separate systematics were added in quadrature, which implicitly assumes no correlation between the separate systematics.

The flux of neutrinos causing a given reaction type in SNO was calculated by multiplying the number of events detected by a normalisation factor:

$$\Phi(X) = \mathcal{N}(X)F(X), \tag{6.18}$$

Appendix I explains how the factor $\mathcal{N}(X)$ was obtained. This factor was also affected by systematic shifts that change the number of simulated events, $E(X)$, passing event selection criteria. The normalisation of the PDFs used in the MXF code gives the predicted number of events, explicitly accounting for corrections due to *sacrifice*, *aborted simulation events*, *energy-dependent fiducial volume uncertainty* and *radiative corrections*.[§] The uncertainty, $\delta_{\mathcal{N}}^{\psi}(X)$ on the normalisation factor for signal X due to systematic shift, ψ, is:

$$\delta_{\mathcal{N}}^{\psi}(X) = \mathcal{N}(X)\left(\frac{E^{\psi}(X) - E^n(X)}{E^n(X)}\right). \tag{6.19}$$

There is an additional contribution to $\delta_{\mathcal{N}}^{\psi}(NC)$ due to uncertainty in the neutron capture efficiency. The uncertainties on $\mathcal{N}(X)$ due to separate systematics were added in quadrature.

6.6 Maximum Likelihood Formalism

To allow for distortions of the electron neutrino energy spectrum, the flux in each T_{eff} bin for the CC and ES signals was treated as a separate fit parameter. The NC flux was treated as a single variable since the energy distribution for NC bears no information on neutrino energy and is simply the energy measured from neutron capture. The number of EN events was also treated as a single fit parameter, whilst the number of internal γ events was fixed in the fit. For the analysis presented in this chapter, the likelihood for each event was:

[§]Some energy-independent corrections were also required as detailed in Appendix I.

$$\begin{aligned}\mathcal{L} &= \sum_{i=\mathrm{CC,ES}} \left(\sum_{k=1}^{N_\mathrm{s}} \left(n_{i_k}.P(i_k/T_{eff}, \beta_{14}, R^3).P(i_k/\cos\theta_\odot/R^3) \right) \right) \\ &+ \sum_{i=\mathrm{NC,EN,int}\gamma} \left(n_i.P(i/T_{eff}, \beta_{14}, R^3).P(i/\cos\theta_\odot/R^3) \right) \end{aligned} \tag{6.20}$$

where n_{i_k} is the flux of signal i in spectral bin k, N_s is the number of spectral bins for the CC and ES distributions, and $P(i_k/X)$ is the probability that an event belongs to signal i, bin k, given its value for observable X.

For each spectral bin, separate PDFs were created for all observables, composed only from simulated events in the given range of T_{eff}. Correlations with T_{eff} for the CC and ES signals, such as that of $\cos\theta_\odot$ are preserved in this way. Other correlations between observable distributions were taken into account with three dimensional PDFs in T_{eff}, β_{14} and R^3 and PDFs for $\cos\theta_\odot$ which were conditional on the value of R^3 (this factorisation of the probability distributions was expressed in equation 5.23).

The validity and impact of the neutron energy information was tested through a separate fit to the salt data in which the number of NC and EN events in each separate T_{eff} bin was also treated as a separate fit parameter. The results of this fit are presented in appendix J and were found to be in very good agreement with the predicted neutron distributions.

6.6.1 Fitting for Total Event Numbers

The total number of CC events obtained from an energy-unconstrained fit is also of great interest as this allows the calculation of a model independent value for the CC to NC flux ratio. Whilst the total number of CC events can be easily obtained by summing the contributions from all spectral bins, it is not trivial to calculate the statistical error on this value as correlations between the fitted fluxes in separate energy bins must be taken into account. This can be calculated using information from the covariance matrix between the fitted event parameters, but the required information can also be obtained from the likelihood fit through a simple change of variables as explained below.

Additional functionality was introduced into the MXF code, such that the total number of CC events replaced the number in the final energy bin, n_{CC_f}, as a fit parameter. A simple conversion was then performed inside the likelihood function to obtain n_{CC_f}:

$$n_{CC_f} = N(\mathrm{CC}) - \sum_{i=1}^{N_s-1} n_{CC_i} \tag{6.21}$$

where $N(\mathrm{CC})$ is the total number of CC events, which is now a parameter of the fit and N_s is the number of spectrum bins. The fitted uncertainty on $N(\mathrm{CC})$ will naturally

include the effects of correlations between the spectrum bins. The parameter $N(\mathrm{ES})$ was also treated as a fit parameter, using the same method. Systematic uncertainties on these parameters were then obtained following the procedure described in section 6.5. Note that this change in fitted variables did not affect the fitted values and uncertainties for the NC and EN signals.

6.7 Results

Figure 6.1 shows the results of the energy unconstrained fit with statistical uncertainties only on the CC and ES spectral bins. The predicted T_{eff} distribution for each signal type is shown scaled by the total fitted flux for that signal. The sum of the fit is shown by a black dashed line which is in good agreement with the data (black stars).

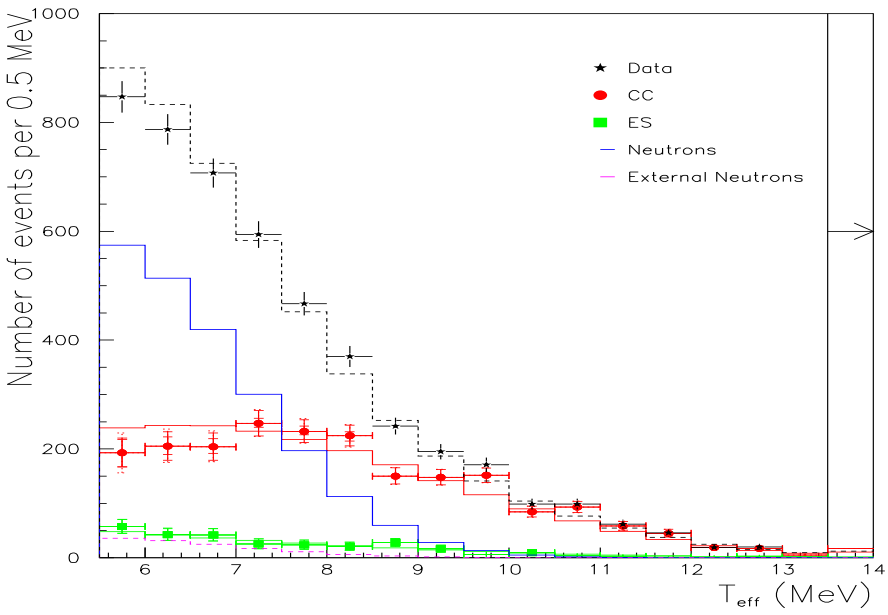


Figure 6.1: Results of the energy unconstrained fit. Only statistical uncertainties are plotted here. The predicted T_{eff} distribution for each signal type is shown scaled by the total fitted flux for that signal. The black dashed line gives the sum of the fit which is in good agreement with the data distribution shown by the black stars. The last energy bin extends to 20.0 MeV.

The total number of events for each type of neutrino interaction (obtained using the method explained in section 6.6.1) are given in table 6.9 with both statistical and systematic uncertainties. The internal neutron background was subtracted from the fitted

Interaction Type	Fitted Number of Events
CC	$2094.6\ ^{+79.7}_{-78.9}$ (stat.) $^{+59.7}_{-60.8}$ (sys.)
ES	$284.0\ ^{+26.9}_{-27.3}$ (stat.) $^{+10.5}_{-8.8}$ (sys.)
NC	$2088.4\ ^{+91.4}_{-90.8}$ (stat.) $^{+62.7}_{-76.2}$ (sys.)

Table 6.9: The total number of events obtained for each signal type from an energy-unconstrained fit to the salt data set. Full statistical and systematic uncertainties are given.

number of neutrons to give the number of NC events, and the additional uncertainty contribution was included in the NC statistical uncertainty.

The CC spectrum is shown in more detail in figure 6.2, with both statistical and systematic uncertainties. The fitted values along with the statistical and combined systematic uncertainties are given in table 6.10. Figure 6.3 shows the contributions of the separate systematic uncertainties to the CC spectrum as a fraction of the CC flux in each energy bin. The dominant systematic uncertainties are associated with T_{eff} and β_{14}.

The combined uncertainties on the measured CC spectrum are $> 15\%$ in the first two spectrum bins, where the predicted MSW distortion is approximately 5%. Since statistical uncertainty dominates this measurement, it will not be possible to achieve the precision required to make a definitive statement about LMA spectral shape distortions using the salt data alone, even if the systematic uncertainties are dramatically reduced because all the available phase II data was analysed here.

6.7.1 Conversion to Fluxes

The normalisation factors calculated for the three measured signals are given in table 6.11, along with the predicted number of events due to ^{8}B and *hep* solar neutrinos for each interaction type from the flux predictions of [21].

There is also an uncertainty on each interaction due to uncertainties in the cross-section values used in the simulation [101]. This is because of a 0.5% uncertainty in the weak axial coupling parameter, g_a [102], and uncertainty in the radiative corrections [103]. The combined contribution of these uncertainties to each flux are also given in table 6.11.

Using equation 6.18, the fluxes for neutrinos causing CC, ES and NC events were calculated to be:

$$\Phi(CC) = 1.62 \pm 0.06(\text{stat.}) \pm 0.07(\text{sys.}) \pm 0.02(\text{theory}) \quad (6.22)$$

$$\Phi(ES) = 2.43 \pm 0.23(\text{stat.}) \pm 0.12(\text{sys.}) \pm 0.01(\text{theory}) \quad (6.23)$$

$$\Phi(NC) = 5.23 \pm 0.23(\text{stat.}) \pm 0.27(\text{sys.}) \pm 0.06(\text{theory}) \quad (6.24)$$

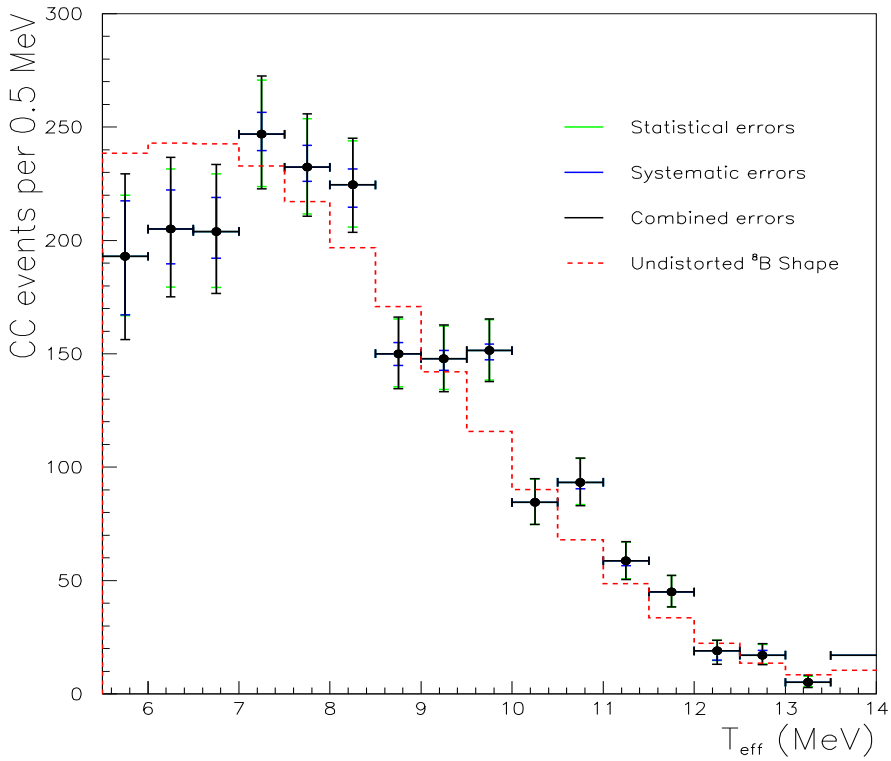


Figure 6.2: CC spectrum obtained from the full salt data set with both statistical and systematic uncertainties. The predicted CC distribution for an undistorted ^{8}B spectrum scaled by the total fitted CC flux is also shown.

Parameter	Fit	Stat.	Sys.	Total
$F(CC)_1$	192.978	$^{+26.952}_{-26.078}$	$^{+24.409}_{-25.820}$	$^{+36.362}_{-36.698}$
$F(CC)_2$	205.007	$^{+26.425}_{-25.583}$	$^{+17.298}_{-15.346}$	$^{+31.583}_{-29.833}$
$F(CC)_3$	203.936	$^{+25.437}_{-24.620}$	$^{+14.979}_{-11.786}$	$^{+29.520}_{-27.295}$
$F(CC)_4$	246.846	$^{+23.773}_{-23.020}$	$^{+9.579}_{-7.169}$	$^{+25.630}_{-24.110}$
$F(CC)_5$	232.298	$^{+21.421}_{-20.688}$	$^{+9.635}_{-6.307}$	$^{+23.488}_{-21.628}$
$F(CC)_6$	224.515	$^{+19.321}_{-18.597}$	$^{+7.017}_{-9.809}$	$^{+20.556}_{-21.025}$
$F(CC)_7$	150.033	$^{+15.329}_{-14.615}$	$^{+4.928}_{-5.134}$	$^{+16.102}_{-15.490}$
$F(CC)_8$	147.864	$^{+14.356}_{-13.649}$	$^{+3.689}_{-5.108}$	$^{+14.822}_{-14.573}$
$F(CC)_9$	151.412	$^{+13.680}_{-12.999}$	$^{+2.892}_{-3.997}$	$^{+13.982}_{-13.600}$
$F(CC)_{10}$	84.483	$^{+10.357}_{-9.642}$	$^{+1.508}_{-1.546}$	$^{+10.466}_{-9.765}$
$F(CC)_{11}$	93.328	$^{+10.554}_{-9.875}$	$^{+1.622}_{-2.842}$	$^{+10.678}_{-10.276}$
$F(CC)_{12}$	58.723	$^{+8.311}_{-7.966}$	$^{+0.984}_{-2.064}$	$^{+8.369}_{-8.229}$
$F(CC)_{13}$	44.965	$^{+7.241}_{-6.579}$	$^{+1.062}_{-0.703}$	$^{+7.318}_{-6.616}$
$F(CC)_{14}$	19.004	$^{+4.694}_{-4.327}$	$^{+0.006}_{-4.089}$	$^{+4.694}_{-5.953}$
$F(CC)_{15}$	17.142	$^{+4.639}_{-3.951}$	$^{+2.107}_{-1.341}$	$^{+5.095}_{-4.172}$
$F(CC)_{16}$	5.148	$^{+2.651}_{-1.972}$	$^{+1.693}_{-1.277}$	$^{+3.146}_{-2.349}$
$F(CC)_{17}$	17.002	$^{+4.461}_{-3.799}$	$^{+0.005}_{-2.828}$	$^{+4.461}_{-4.736}$
$F(ES)_1$	57.508	$^{+13.195}_{-12.275}$	$^{+2.020}_{-2.526}$	$^{+13.349}_{-12.532}$
$F(ES)_2$	42.373	$^{+11.345}_{-10.422}$	$^{+4.085}_{-1.465}$	$^{+12.058}_{-10.524}$
$F(ES)_3$	41.797	$^{+11.559}_{-10.626}$	$^{+2.429}_{-1.236}$	$^{+11.811}_{-10.698}$
$F(ES)_4$	25.248	$^{+9.349}_{-8.403}$	$^{+0.706}_{-2.145}$	$^{+9.376}_{-8.672}$
$F(ES)_5$	23.750	$^{+8.787}_{-7.813}$	$^{+1.972}_{-2.357}$	$^{+9.006}_{-8.161}$
$F(ES)_6$	20.897	$^{+7.883}_{-6.999}$	$^{+1.715}_{-1.264}$	$^{+8.067}_{-7.112}$
$F(ES)_7$	27.971	$^{+7.498}_{-6.661}$	$^{+1.312}_{-3.817}$	$^{+7.612}_{-7.677}$
$F(ES)_8$	16.654	$^{+6.744}_{-5.867}$	$^{+0.690}_{-2.451}$	$^{+6.779}_{-6.358}$
$F(ES)_9$	6.072	$^{+4.966}_{-3.997}$	$^{+0.720}_{-1.624}$	$^{+5.018}_{-4.314}$
$F(ES)_{10}$	9.538	$^{+4.921}_{-4.138}$	$^{+1.202}_{-2.094}$	$^{+5.066}_{-4.637}$
$F(ES)_{11}$	3.845	$^{+4.087}_{-3.135}$	$^{+2.019}_{-1.480}$	$^{+4.558}_{-3.467}$
$F(ES)_{12}$	2.652	$^{+3.101}_{-2.144}$	$^{+1.917}_{-0.868}$	$^{+3.646}_{-2.313}$
$F(ES)_{13}$	0.939	$^{+2.419}_{-1.896}$	$^{+0.511}_{-1.053}$	$^{+2.472}_{-2.169}$
$F(ES)_{14}$	0.000	$^{+2.054}_{-0.022}$	$^{+3.772}_{-0.000}$	$^{+4.295}_{-0.022}$
$F(ES)_{15}$	2.860	$^{+2.414}_{-1.657}$	$^{+1.329}_{-2.111}$	$^{+2.756}_{-2.684}$
$F(ES)_{16}$	1.855	$^{+1.763}_{-1.100}$	$^{+1.266}_{-1.698}$	$^{+2.170}_{-2.023}$
$F(ES)_{17}$	0.000	$^{+1.110}_{-0.002}$	$^{+1.013}_{-0.000}$	$^{+1.503}_{-0.002}$
$F(neutron)$	2226.396	$^{+86.115}_{-85.091}$	$^{+62.783}_{-76.405}$	$^{+106.572}_{-114.360}$
$F(EN)$	133.461	$^{+42.805}_{-41.837}$	$^{+33.417}_{-17.090}$	$^{+54.304}_{-45.193}$
$F(in\gamma)$	8.500	$^{+0.000}_{-0.000}$	$^{+2.200}_{-2.200}$	$^{+2.200}_{-2.200}$

Table 6.10: Fitted fluxes with their statistical, systematic and combined errors. Separate results for each of the 17 CC and ES bins are given with bin 1 being the lowest in energy (5.5–6.0MeV). The fitted neutron flux contains a contribution from the internal neutron background as well as the NC flux. The internal gamma background was fixed to 8.5 events in the fit.

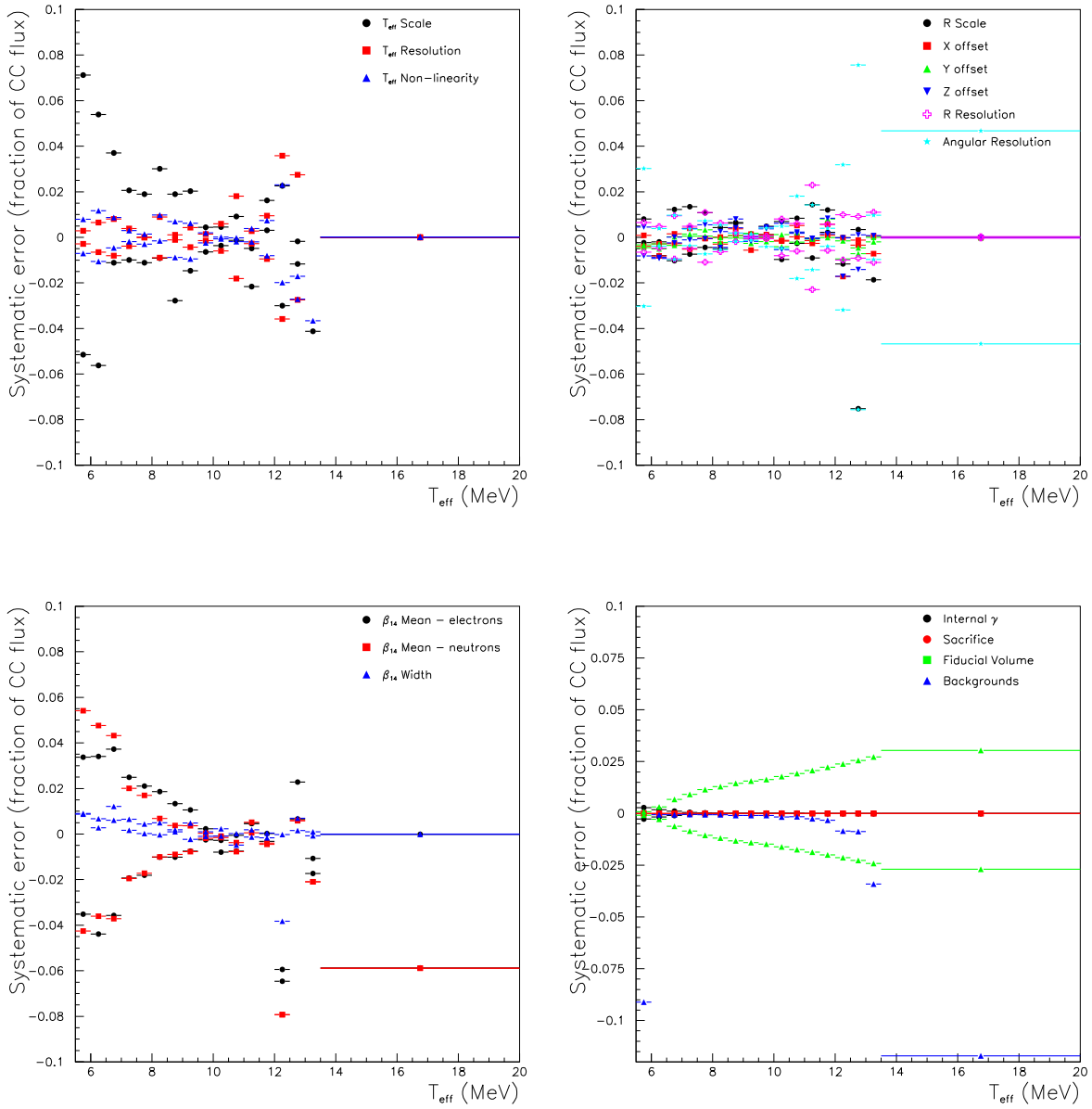


Figure 6.3: Contributions of systematic uncertainties to the CC energy spectrum given as a fraction of the flux in each energy bin. The separate panels show T_{eff} systematics, reconstruction systematics, β_{14} systematics, and PDF weighting or background systematics.

Interaction Type	$E(X)$		$\mathcal{N}(X)$	$\Delta(\sigma)$
	^{8}B	*hep*		
CC	7506^{+255}_{-248}	$29.85^{+1.01}_{-0.98}$	$773.3^{+26.2}_{-25.4}$	$\pm$ 1.2%
ES	678^{+27}_{-26}	$2.20^{+0.09}_{-0.08}$	$8564.3^{+343.2}_{-317.9}$	$\pm$ 0.5%
NC	2330^{+99}_{-95}		$2501.7^{+111.0}_{-106.1}$	$\pm$ 1.1%

Table 6.11: Expected numbers of events, $E(X)$, and normalisation factors, $\mathcal{N}(X)$, for conversion to fluxes (as specified in equation 6.18) including systematic uncertainties (see equation 6.19). Cross-section uncertainties, $\Delta(\sigma)$ are also given for each interaction type (taken from Ref. [31])

	CC	ES	NC+INB	EN
CC	5648.7427	-243.3250	-3199.7639	-276.1071
ES	-243.3250	724.1163	-174.6347	-20.3017
NC+INB	-3199.7639	-174.6347	7088.7314	-1363.0956
EN	-276.1071	-20.3017	-1363.0956	1795.7030

Table 6.12: Parameters of the covariance matrix between the fitted total fluxes.

all in units of $10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The NC flux is in good agreement with the solar model prediction of

$$\Phi_{\mathrm{SSM}} = 5.83 \pm 1.34(\mathrm{theory})$$

(which includes the *hep* contribution), whilst there is a significant reduction for the CC and ES fluxes.

The ratio of the CC to NC fluxes was calculated to be:

$$\frac{\Phi(CC)}{\Phi(NC)} = 0.310 \pm 0.021(\mathrm{stat.}) \pm 0.024(\mathrm{sys.}) \tag{6.25}$$

which is in good agreement with predictions for this ratio for the LMA MSW scenario. This measurement is also very close to the previously published ratio [31] given in equation 1.19 which was obtained from a sub-set of the data analysed here.

The covariance matrix between the fitted CC, ES, (NC+INB) and EN parameters is given in table 6.12. This covariance matrix was assumed to described the correlations for both statistical and systematic uncertainties. All reconstruction uncertainties, with the exception of the energy-dependent fiducial volume uncertainty were assumed to have the same effect on both CC and NC events, which would therefore cancel in the ratio. Theoretical cross-section uncertainties were also assumed to cancel.

Data Set	Livetime (days)	Candidate Neutrino Events
Day	176.511	2147
Night	214.921	2603
Total	391.432	4750

Table 6.13: Corrected livetimes and numbers of candidate neutrino events for the day, night and total salt data sets.

6.8 Day and Night CC Spectra

To specifically test for the "day-night effect" (described in section 1.11.1) the MXF code was also used to extract separate *day* and *night* spectra.

Day was defined as any time when the Sun was $\leq 90°$ from the zenith at the location of the SNO experiment, whilst *night* was defined as any time when the Sun was $> 90°$ from the zenith. This information was stored in the data stream so that events could be easily assigned to the *day* or *night* data set. As more work was carried out on the detector in the day time, for both calibrations and repairs, the amount of neutrino data collected at night was greater. The separate day and night livetimes and the number of candidate neutrino events in each data set are given in table 6.13.

For the purposes of this thesis, no constraint was placed on the day-night asymmetry of the NC flux[¶] and the spectrum analysis described in this chapter was simply repeated for two separate data sets consisting only of events arriving in the day and night respectively. A separate day-night analysis, in which the asymmetry in the NC flux was constrained to be zero was performed by the SNO collaboration and an asymmetry of [104]:

$$A_{\rm DN}^{\rm CC} = -0.034 \pm 0.069 \tag{6.26}$$

was measured for the CC flux. An asymmetry of:

$$A_{\rm DN}^{\rm ES} = -0.160 \pm 0.199 \tag{6.27}$$

was obtained for the ES flux. Both of these measurements are consistent with zero asymmetry.

The CC spectra obtained from fits with MXF, presented as the number of events in each energy bin for day and night, are shown in figure 6.4. A chi-squared test was conducted for the two extracted CC spectra, weighted by their respective livetimes. For

[¶]Regeneration effects in the Earth would only affect the measured CC and ES fluxes as the NC flux is sensitive to all active flavours. The functionality to constrain asymmetries had not been incorporated in the MXF code at the time of writing.

17 degrees of freedom, the calculated chi-squared between the two spectra was:

$$\chi^2 = 20.455. \tag{6.28}$$

This relates to 1.2 per degree of freedom which has a probability of roughly 25% [86]. Therefore, there is no evidence for a day-night effect in the shape of the CC energy spectra.

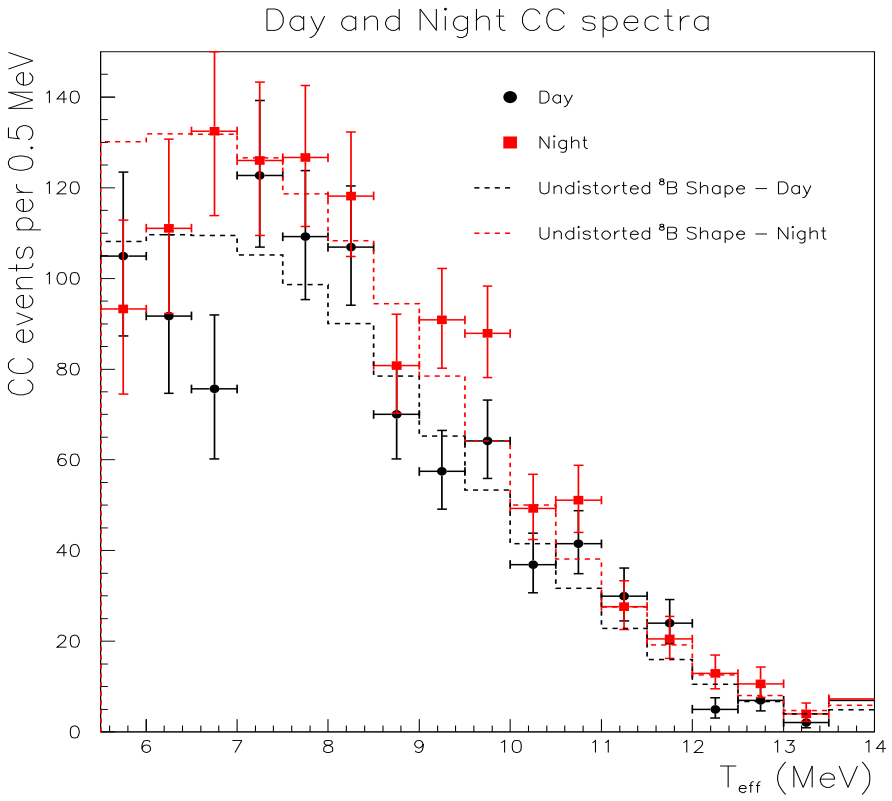


Figure 6.4: Extracted day and night CC energy spectra. The dashed lines show the undistorted ^{8}B CC shape prediction, scaled by the total number of fitted CC events for day and night separately.

These separate day and night CC spectra, were used in a global χ^2 minimisation fit similar to that described in section 1.7.2. The systematic uncertainties evaluated in the total CC spectrum were used to provide uncertainties on the separate day and night spectra in this fit, but the asymmetry in these systematic contributions was obtained from the asymmetry-constrained analysis [104]. The global fit also included results from the radiochemical experiments, SuperK, KamLAND[‖] and the SNO D_2O phase. The contours

[‖]Only the first published KamLAND data [32] was used in these fits. New results were released in June 2004 [105] after this analysis was performed, which further reduce the contour space.

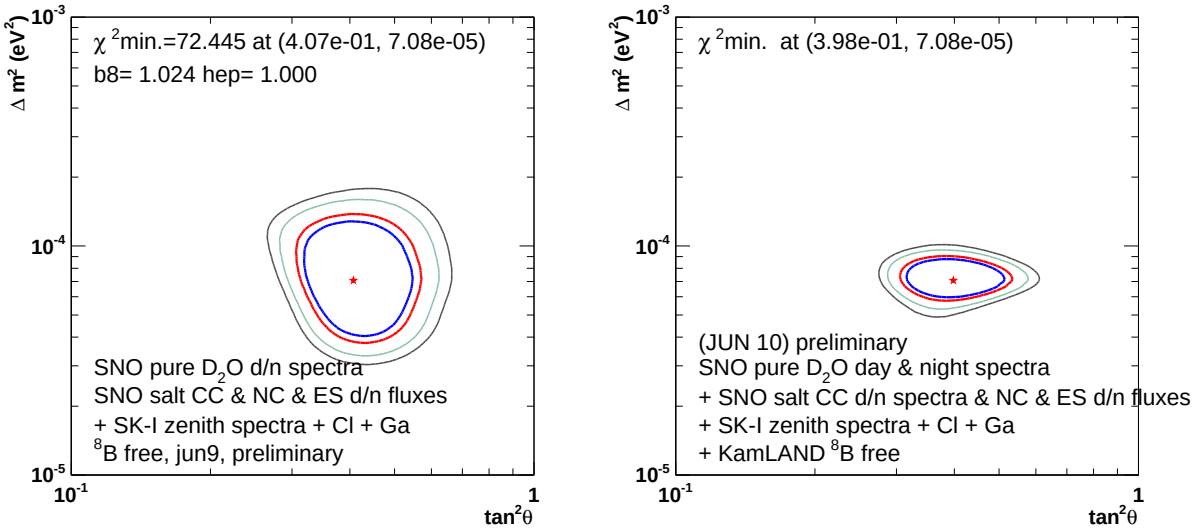


Figure 6.5: Results of a global fit combining day and night CC spectra and ES and NC fluxes from the salt phase with SNO D_2O data, results from radiochemical experiments and the SuperK zenith angle spectra. In the right hand figure KamLAND data [32] has also been included. These are preliminary plots obtained from [106].

for χ^2 confidence levels relating to 90%, 95%, 99% and 99.73% probability (χ^2 = 4.61, 5.99, 9.21 and 11.83 respectively) from this analysis are shown in figure 6.5. In order to make comparisons to the plots shown in figure 1.3, separate plots are shown with and without data from the KamLAND experiment.

6.9 Summary

In this chapter systematic effects on observable parameter distributions have been discussed in detail and included in an energy-unconstrained likelihood fit to the salt data set. The charged-current spectrum measured is consistent with the predicted spectrum shape for no MSW distortions. However, the statistically dominated uncertainties in the measured spectrum are too large to directly distinguish between different MSW mixing scenarios.

Total fluxes for the three types of neutrino interaction obtained from this fit were also presented. The NC flux measurement, which is independent of assumptions of the electron neutrino energy spectrum, is in good agreement with the predicted flux of solar neutrinos, and the CC/NC ratio is in good agreement with the LMA MSW solution.

Spectra were also obtained separately for day and night, which were found to be consistent at the 25% level. These spectra were combined with the flux measurements and other available solar and reactor data to improve the limits on allowed parameter

space for oscillations between the ν_1 and ν_2 mass states.

6.10 Addendum

For completeness, a few improvements on the analysis presented in this chapter are detailed below. These details arose after the analysis presented in this chapter was completed and the effect of these corrections on the obtained CC energy spectrum is small with respect to the magnitude of statistical uncertainty.

The function used to correct for resolution systematics, given in equation 6.2, does not correctly account for the existing detector resolution. The effect of this error is small but for completeness the correct function is:

$$T'_{eff} = T_{eff} + g(\omega_{\mathrm{Tres}}) \tag{6.29}$$

where $g(\delta_{\mathrm{Tres}})$ is a value randomly selected from a Gaussian distribution with zero mean and width

$$\omega_{\mathrm{Tres}} = \sigma_{\mathrm{E}} \sqrt{((1+\delta_{\mathrm{Tres}})^2 - 1)} \tag{6.30}$$

where σ_{E} is the nominal detector resolution which was given previously in equation 5.17. Similarly, the vertex resolution uncertainty given in section 6.2.4 should be defined with the function

$$\omega_R = \sqrt{((1+0.015)^2 - 1)} \times 16\,\mathrm{cm} \tag{6.31}$$

which assumes a nominal vertex resolution of 16 cm, with a 1.5% uncertainty.

The effect of the radial bias in energy is actually accounted for by the RSP processor, and therefore should not be applied as a correction. Instead, the magnitude of this correction should be taken as a systematic uncertainty to account for possible inaccuracies in the RSP calculation of this effect. Again, the effect of this correction on the fitted central values is small with respect to the statistical uncertainty in the spectrum.

An additional improvement to the analysis presented here would be the use of energy-dependent mean values for β_{14} for each signal. Parameterisations of the mean β_{14} value for electron and neutron-type events are provided in [57].

Chapter 7

A Different Approach to Systematic Uncertainties

7.1 Motivation

The analysis presented in the previous chapter allowed for a number of systematic uncertainties in the fitted CC spectrum. The dominant systematic uncertainties were those associated with energy and isotropy. The effect of these uncertainties was previously estimated by performing multiple fits to the data with different PDFs. This approach assumed that all the systematics were uncorrelated and that the likelihood function was normally distributed for each systematic parameter; only the $\pm 1\sigma$ deviations were tested.

A more rigorous *smearing integration method* tests many values for each systematic. For a series of fits to the same data set, shifts are applied to all PDF observables simultaneously. The magnitude of each systematic shift is randomly selected from an assumed probability distribution. The width of the distribution of fitted values, then gives the uncertainty on each variable. This technique is extremely computationally intensive and correlations between the separate systematic uncertainties could only be included if determined externally to the fitting procedure.

An alternative approach is to treat systematic changes in observable distributions as parameters of the likelihood fit. This allows for the best estimate of these systematic parameters based on the maximum likelihood and a formal treatment of correlations between the uncertainties. Estimates for the magnitude of these systematics, obtained from studies of calibration data, can also be input to the likelihood fit as constraints on the fitted systematic parameters. A simple example of the benefits of this approach is given in the following section.

7.1.1 Why Treat Systematics as Fit Parameters?

This section details a simple example which demonstrates how treating systematic uncertainties on observable distributions as fit parameters can result in an overall reduction of uncertainties on the fitted fluxes. Treating systematic uncertainties as extra degrees of freedom in likelihood space essentially accounts for correlations between all the fit parameters and produces combined statistical and systematic uncertainties.

A simulated data set was generated consisting of 1000 events sampled in equal proportions from two Gaussian distributions, A and B, in x (figure 7.1). Binned PDFs were created from a million generated events for each distribution. MXF was used to extract the number of events of type A and type B from the data set,* minimising over a 2 dimensional likelihood space. The result was:

$$\begin{aligned} A &= 491 \pm 28 \\ B &= 507 \pm 28 \end{aligned}$$

Now assume there was a 5% systematic uncertainty in the scale of parameter x between the data set and the PDF used. In other words, the predicted distribution might incorrectly be scaled by 5% in either direction with respect to the data.† Adopting the approach used in the previous chapter, the effect of this uncertainty was determined by recreating the PDFs, scaling observable x up (down) by 5% as shown by the red (blue) plots in figure 7.1. The fit was repeated with the 1σ PDFs to give the following fit results, with the difference from the original fit given in brackets. The fit with PDFs scaled up by 1σ gave:

$$\begin{aligned} A &= 413(-78) \\ B &= 586(+79). \end{aligned}$$

The fit with PDFs scaled down by 1σ gave:

$$\begin{aligned} A &= 574(+83) \\ B &= 425(-82). \end{aligned}$$

These systematic uncertainties were then combined in quadrature with the statistical uncertainties on the original fit to give the following total uncertainties on the fitted fluxes:

$$\begin{aligned} A &= 491^{+88}_{-83} \\ B &= 507^{+84}_{-87} \end{aligned}$$

*The fit was performed for values $x > 0.0$, which excluded two events from the data sample.

†Here a large uncertainty is used for illustration.

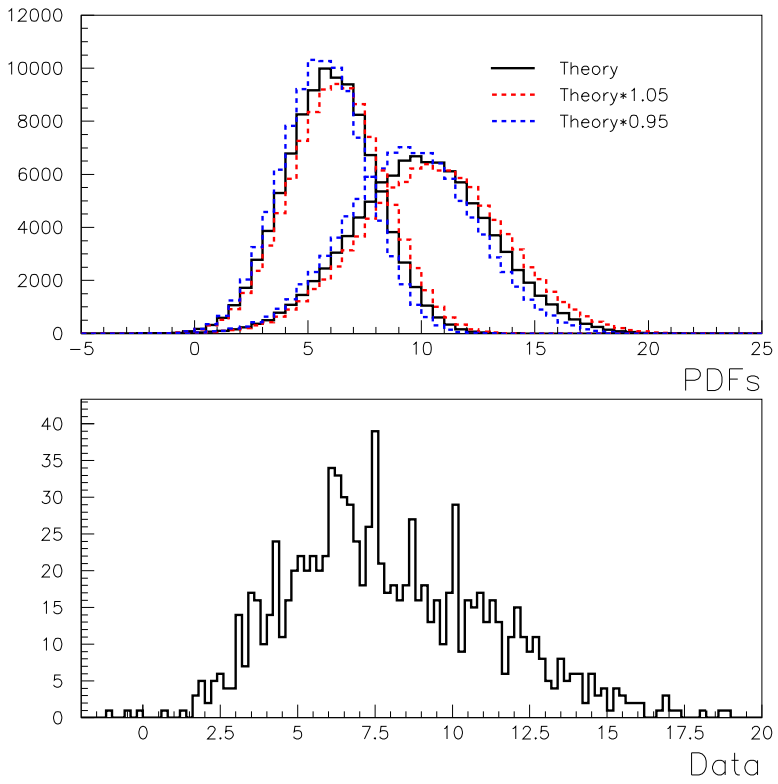


Figure 7.1: PDFs (upper panel) and data set (lower panel) for the simple example described in section 7.1.1. The top panel shows the default PDFs in black solid lines and PDFs scaled by ± 5% in red and blue dashed lines.

For the alternative approach, the maximum likelihood value was obtained by varying the systematic scale as well as the magnitude of signal A and B. The same 1σ uncertainty of 5% was used to constrain the systematic parameter in the fit (see 5.1.1 for details). The result from minimisation over three dimensional likelihood space was:[‡]

$$\begin{aligned} A &= 487^{+60}_{-64} \\ B &= 511^{+66}_{-40} \\ scale &= 0.0025^{+0.029}_{-0.014}. \end{aligned}$$

These results are in good agreement with the original fit. The scale fitted on observable x is consistent with zero, as expected, but the combined statistical and systematic uncertainties are noticeably smaller.

7.2 Dominant Systematic Uncertainties

Additional fit parameters increase the complexity of the likelihood fit and hence the time required to perform the fit. The analysis presented in the previous chapter showed that only a few of the systematics on observable parameters made a significant contribution to the uncertainties on the fitted fluxes. Therefore, for the purpose of this thesis, only these dominant systematics were varied in the likelihood fit. This section summarises the dominant observable systematic uncertainties in the context of an energy-constrained fit.

7.2.1 Energy Uncertainties

As discussed in 6.2.1, there are three components to systematic uncertainty in the T_{eff} distributions. Potential non-linear effects were assumed to modify the T_{eff} distribution in a similar manner to the energy scale systematic. These uncertainties were accommodated with two additional variables in the likelihood fit: a scale parameter allowed the predicted T_{eff} distribution to scale with respect to the data T_{eff} value, whilst an offset parameter essentially changed the point at which the effect of this scale was zero. The scale parameter was constrained in the likelihood fit by $\pm$ **1.15%** [88], which included 0.25% due to possible non-linear effects,[§] whilst the offset was constrained by $1.15\% \times 5.0\,\mathrm{MeV} = \pm$ **0.0575 MeV**.

The treatment of energy resolution uncertainty in chapter 6 showed that the effect of this systematic was considerably smaller than that of energy scale, and therefore, was not included as a fit parameter.

[‡]PDF smoothing as described in 7.4.1 was employed to obtain these results.

[§]This value of 0.25% was adopted for previous analyses of both D_2O and salt data.

7.2.2 Reconstruction Uncertainties

Uncertainties in reconstructed vertex position were discussed in section 6.2.4. The dominant reconstruction uncertainty is due to a possible systematic in the reconstructed event position that scales with radius from the center of the detector, R. This systematic was included in the likelihood fit as a scale applied to R and not R^3, constrained with a limit of $\sigma = \pm\mathbf{1.0\%}$ [95].

7.2.3 Uncertainties in Isotropy Parameters

Instead of using the energy-dependent limits on β_{14} systematics discussed in the previous chapter, a single value was used to limit the size of a systematic error in the mean and width of the predicted isotropy distributions. Above the detector energy threshold, integral values of 0.85% uncertainty on the mean and 0.94% uncertainty on the width of the β_{14} distribution were estimated based on studies of ^{16}N and ^{252}Cf calibration data [57]. The proposed systematic limits for mean and width were translated into approximate limits on a scale and offset of the β_{14} values as follows:

Equation 7.1 shows how an observable, x, would be shifted given a fractional systematic shift of $d\bar{x}$ on the mean, $\bar{x}$, whilst equation 7.2 shows how the observable would be shifted given a systematic $d\omega$ on the width ω of the distribution.

$$x' = x \pm \bar{x}.d\bar{x} \tag{7.1}$$

$$x' = \bar{x} + (x - \bar{x}) \times (1 \pm d\omega) \tag{7.2}$$

Applying both systematics simultaneously to x in the positive direction results in a new value, x' given in equation 7.3

$$x' = \bar{x} + (x + \bar{x}.d\bar{x} - \bar{x}) \times (1 + d\omega) \tag{7.3}$$

Assuming no correlation between the two systematic effects, this can be approximated as a scale, α and offset β:

$$\begin{aligned} x' &= x.\alpha + \beta \\ \alpha &= 1 + d\omega \\ \beta &= \bar{x}.(d\bar{x} + d\bar{x}d\omega - d\omega) \end{aligned} \tag{7.4}$$

These values for α and β were used as 1σ Gaussian constraints on the systematic uncertainty.

To test that this formalism was sufficient, 100 simulated data sets of 10^6 events each were generated from a Gaussian distribution in x with $\bar{x} = 0.48$ and $\omega = 0.133$, the mean

Observable	Systematic Parameter	Constraint
T_{eff}	Scale	$\pm 1.15\%$
T_{eff}	Offset	± 0.0575 MeV
R	Scale	$\pm 1.0\%$
β_{14}	Scale	$\pm 0.94\%$
β_{14}	Offset	± 0.00039

Table 7.1: The dominant uncertainties on observable distributions that are treated as parameters of the likelihood fit. Constraints applied to each parameter in the fit are also given.

and width for the CC β_{14} distribution. For each of the 100 data sets, each event was shifted by a value of $d\bar{x}$ selected from a Gaussian distribution with zero mean and width 0.0085, and a value of $d\omega$ selected from a Gaussian with zero mean and width 0.0095, following the prescription of equation 7.3.

These values of $d\bar{x}$ and $d\omega$ were translated into $\alpha = 1.0094$ and $\beta = -0.0008201 \times 0.48$. A second set of 100 data sets were generated with the same statistics but, this time, each event was shifted by a value of α selected from a Gaussian distribution with width 1.0094 and β with a value selected from a Gaussian distribution of width -0.00039, following the prescription given in equation 7.4.

The distribution of $\bar{x}$ for the simulated data sets was found to be slightly broader than expected in both cases, which could be due to correlations between the two systematic uncertainties. The prescription of equation 7.4 resulted in only slightly broader distributions of $\bar{x}$ and ω.

The limits on α and β derived in this manner were used as 1σ constraints on a variable scale and offset of β_{14} in the likelihood fit. The mean and width of the β_{14} distribution is different for each signal (see table 6.1), resulting in a larger change in β_{14} value for electron-type events than neutron-type events. However, for the purposes of this thesis, the uncertainty on the CC β_{14} distribution was used, which is more conservative and slightly overestimates the uncertainty on the NC distribution. The scale on β_{14} was constrained by $\sigma = \pm$ **0.0094**, whilst the width parameter was constrained by $\sigma = \pm$ **0.00039**.

7.2.4 Observable Systematics as Fit Parameters - Summary

Table 7.1 summarises the 1σ limits on the dominant uncertainties on observable parameters that were treated as parameters of the likelihood fit.

7.3 Incorporating Systematics in MXF

This section details the method used to include parameters describing a scale or offset between data and simulation for a given observable in the likelihood fit. These systematic parameters were also constrained in the fit by incorporating the best estimates from independent studies as a Gaussian constraint within the likelihood function.

7.3.1 Systematic Scale

A difference in scale, α, on an observable parameter can be included in the fit by scaling that observable for each event contributing to the PDF by $(1+\alpha)$. However, scaling each Monte Carlo value and recreating the PDF "on the fly" is extremely time consuming. Instead, the systematic scale was included by scaling each *data value* by a factor $\left(\frac{1}{1+\alpha}\right)$. This effectively changes the data range for the scaled observable, x from $(x_{max} - x_{min})$ to $((1+\alpha)x_{max} - (1+\alpha)x_{min})$. As mentioned earlier, PDFs must be normalised over the same range as the data so the PDF must be renormalised over the new data range. This is described in more detail in 7.4.2. As the data maximum can increase (or the minimum can decrease), it is necessary to store information on an observable over a greater range than the data. For this purpose, additional bins were stored at each end of a PDF called *overflow bins*, that only contributed to the normalisation when systematics were non-zero in the fit. If a data value was shifted beyond this range, the log likelihood was set to zero for that event.

An additional correction to the normalisation was required because the systematic scale also changes the width of each bin (or the numerical integration step for analytic PDFs). Consider the binned PDF case: the normalised amplitude, $A(i)$, for each bin, i, is given by equation 7.5:

$$A(i) = \frac{n(i)}{T \times w} \tag{7.5}$$

where $n(i)$ is the number of Monte Carlo events falling in that bin, T is the total number of events in the data range for that PDF and w is the bin width. Renormalisation accounts for the change in T but, to account for the change in bin width the probability obtained from a scaled PDF was multiplied by the factor $\frac{1}{(1+\alpha)}$.

7.3.2 Systematic Offset

Similarly, a systematic offset, β, between data and simulation in a given observable was taken into account by shifting each data value by $-\beta$ in that observable. The correction for bin width was not necessary in this case but renormalisation of the PDFs was required.

7.4 Working with Discrete PDFs

Applying systematic variables in the fit can cause discrete changes in the likelihood function when binned PDFs are used. The blue triangles in figure 7.2 show how the likelihood space looks for the simple scenario described in section 7.1.1. The bumps are caused by the way data events are moved across bin boundaries as the value of the systematic scale is changed.

Uneven likelihood functions were found to be problematic to minimise. In many cases, parameters at local minima instead of the true maximum likelihood were fitted[¶] and the estimated uncertainties on systematic parameters were unreasonably small.

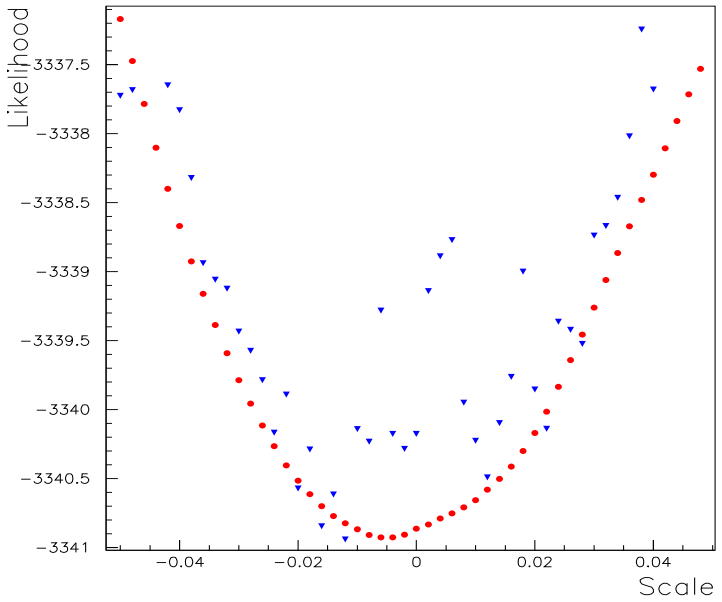


Figure 7.2: Likelihood space in the scale parameter for the simple example described in 7.1.1. Blue triangles are for a fit with discrete binned PDFs, and red circles are for a fit with PDFs smoothed by a linear interpolation.

7.4.1 Smoothing PDFs

To overcome the problems associated with bumpy likelihood space, the binned PDFs were smoothed to provide continuous probability distributions. Care was taken to ensure that the smoothing procedure did not introduce additional systematics to the PDF shape.

[¶]Minimisation routines work on the negative of the likelihood

A polynomial smoothing algorithm was implemented in the MXF code to smooth PDFs "on the fly". A linear interpolation between adjacent bins effectively smoothed the likelihood space, but this first order smoothing distorted the shapes of some observable distributions and was found to cause biases in the results of the fit. Higher orders of polynomial smoothing were found to improve the shape of the smoothed distribution.

For nth order polynomial smoothing, the function is fit to the $n+1$ central bin values lying closest to x. However, even this can still lead to small discontinuities in the smoothed PDFs at bin boundaries. For example, if $n = 2$, then 3 bin values are used for the smoothing. If the data value lies in bin α near the lower bin boundary, then the bins $(\alpha-1)$, α and $(\alpha+1)$ will be used by the smoothing algorithm. However if the data value lies in bin $(\alpha-1)$ near the upper bin boundary, bins $(\alpha-2)$, $(\alpha-1)$ and α will be used. To avoid such discontinuities, odd values of n were used for smoothing.

Multi-dimensional PDFs can be smoothed in two dimensions through two successive, one-dimensional interpolations. For nth order smoothing in both dimensions (x and y), an $(n+1) \times (n+1)$ grid centered about the desired point (x_1, y_1) was selected. $(n+1)$ 1-dimensional interpolations were carried out in the y direction to obtain probabilities at the points $(x(j), y_1)$ for $j = 1, ...(n+1)$, followed by a single interpolation in the x direction.

If the central values of PDF bins are not equally spaced, the smoothing procedure can bias the PDF shape. Therefore, equal sized bins were chosen for all observable distributions that required smoothing.

nth order smoothing in x requires $\frac{n+1}{2}$ central bin values on either side of X. This can lead to problems when X is close to the data limit for x, as it may be necessary to extrapolate beyond the last bin in a PDF. Overflow bins solve this problem, providing information on the distribution of the observable beyond the data limits. A number of equally sized overflow bins were created for interpolation beyond the regions of a PDF sampled by the likelihood function.

Smoothing will naturally change the normalisation of a binned PDF so smoothed PDFs were renormalised within the MXF code.

7.4.2 Renormalisation of PDFs

The normalisation of binned PDFs was calculated at the time of creation by summing the weighting of events contributing to each PDF. For analytic PDFs, the normalisation was determined *via* numerical integration. For smoothed PDFs the normalisation over the specified data range was also determined *via* numerical integration.

If a systematic shift is applied to observable x, the data range will change and the normalisation must be recalculated for each different systematic value tested by the likeli-

hood function. A complete numerical integration over the whole PDFis time consuming, especially for multi-dimensional PDFs. Instead, the approach of MXF was to numerical integrate over the range that is either added or subtracted at the end of the PDF. This is illustrated in figure 7.3.

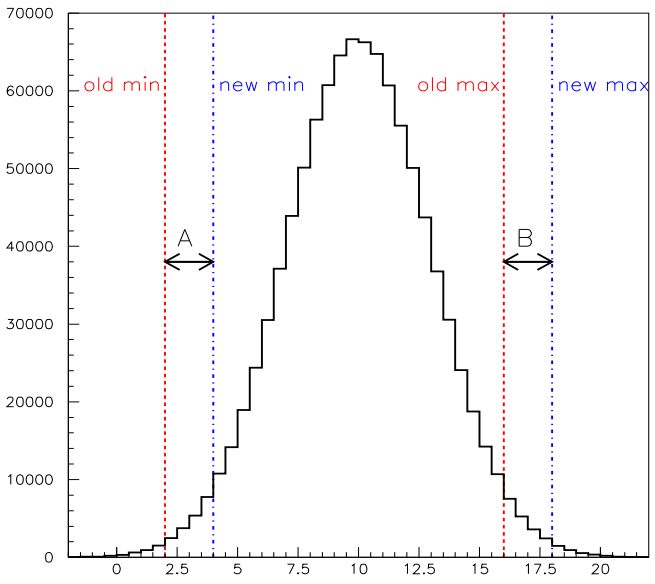


Figure 7.3: Illustration of PDF renormalisation when systematics are replied. The numerical integral over A is subtracted from the original normalisation and the numerical integral over B is added to obtain the new normalisation consistent with systematics which shift the data and minimum as shown.

A similar procedure was adopted for PDFs with systematics applied in 2 dimensions, with the upper and lower ends of the PDF in each observable treated separately.

7.5 Further Complications to the Fitting Procedure

The PDF smoothing and renormalisation procedure becomes more complicated for multi-dimensional PDFs, and to date, has only been applied to one and two-dimensional distributions. To achieve the desired level of accuracy, a large number of steps must be used for the numerical integration, which makes this process much slower for distributions in more than two dimensions, even when the approach described in section 7.4.2 is used.

A more serious problem arises for conditional PDFs, which are normalised separately for each bin in the conditional parameter. Thus, discrete steps remain in the distribution

Signal	A	B	Correction factor
CC	$2233.36^{+69.5}_{-68.6}$	$2209.66^{+69.7}_{-68.8}$	+1.07%
ES	$279.16^{+28.2}_{-27.3}$	$283.47^{+28.4}_{-27.4}$	-1.52%
NC	$2105.49^{+76.6}_{-75.5}$	$2111.98^{+78.3}_{-77.2}$	-0.31%
EN	$120.7^{+42.5}_{-41.6}$	$135.6^{+43.6}_{-42.8}$	-10.99%

Table 7.2: Results for fits to salt data using different PDF factorisations (A = bias-free factorisation of equation 5.22, B = factorisation adopted for the analysis presented in this chapter), and the correction factors derived from them.

of the conditional parameter that cannot be removed by smoothing algorithms. Because these discrete steps are intrinsic to the distribution, it is difficult to treat systematics on the conditional parameter as variables in the likelihood fit.

To overcome these problems, the factorisation of probabilities given in equation 5.21 was adopted for the analysis presented in this chapter. This required only one multi-dimensional PDF to account for the correlation between T_{eff} and β_{14}. The results from such a fit are subject to a small bias of around 1% on the fitted fluxes. This was corrected after the fit as described in section 7.6.

For the energy-unconstrained fit described in the previous chapter, the number of events in each T_{eff} bin was treated as a separate fit parameter. For such a fit the distribution in T_{eff} is naturally discrete and varying systematics on this observable will also result in uneven likelihood space. This is not a problem for the energy-constrained fit described in this chapter or for fits for fluxes in separate *neutrino* energy bins as described in chapter 8.

7.6 Correcting Biases Due to PDF Factorisation

For the purpose of this thesis, a data-based correction was derived to correct the fitted fluxes for biases introduced by the PDF factorisation used. The energy-constrained fit was performed on the salt data using the PDF factorisation specified in equation 5.22. The results of this fit, which did not include any systematic parameters, are given in column 'A' of table 7.2. The results of a fit using the chosen PDF factorisation, again without any systematic parameters, are given in column 'B'. The percentage correction factors given in this table were obtained from a direct comparison of the two fit results.

X	$F(X)$	Mean $Bias(X)$	$\Delta(X)$	$\bar{\sigma}(X)$
CC	2004.6 ± 6.9	-0.31 ±0.06	-0.011 ±0.002	68.556 ±0.096
ES	317.3 ± 2.9	0.30 ±0.07	0.030 ±0.007	29.367 ±0.067
NC	2168.9 ± 7.8	-0.25 ±0.07	-0.009 ±0.003	78.375 ±0.091
EN	179.0 ± 4.5	0.68 ±0.09	0.206 ±0.025	44.914 ±0.077

Table 7.3: Results for fits to 100 artificial data sets with PDFs smoothed by a third order polynomial.

7.7 Verification of the Technique

To study possible biases introduced to the fitting procedure through the smoothing of PDFs, tests were carried out on the 100 simulated data sets described in 5.7. The specific likelihood function given in equation 5.20 with PDFs factorised as given in equation 5.21 was used throughout. The PDFs were created using the binning specified in table 5.1, with the exception of T_{eff}, where 29 equal sized bins were used. Distributions in T_{eff}, β_{14} and R^3 were smoothed by a third order polynomial. The results of fits to the artificial data sets are shown in table 7.3.

No bias correction has been applied to these results. They can be compared to those given in table 5.3, which were obtained with the same factorisation of PDFs, but no smoothing. The bias on the external neutron and charged current fluxes appears to be around 1σ larger for the smoothed PDF approach. However, a change of 1σ is relatively modest considering that the fractional biases in the fitted signals are still only about 1%. Thus, the smoothed PDFs can be used as a good approximation for the predicted distributions.

A further test was carried out on the same artificial data sets, with the same PDFs, but with five additional fit parameters relating to the scale and offset type systematic parameters given in table 7.1. The results of these fits are given in table 7.4. The fractional difference $\Delta(X)$ is not a meaningful parameter for the systematic fit parameters, where the expected value is zero in each case. Instead, the consistency of the mean value of the fitted systematic, $F(X)$, with zero can be used.

For most of the systematic fit parameters, the mean values are consistent with zero. However, the radial scale systematic parameter appears to be biased at a level of $\approx 4.4\sigma$. Whilst a small bias is expected in the fitted fluxes due to the factorisation of PDFs used, the magnitude of the bias in the NC and EN signals is larger than previously observed with the same factorisation. The effect could be due to artefacts introduced by smoothing of the coarsely binned radial PDFs. The results of a second set of fits (shown in table 7.5), performed with radial PDFs created in 20 bins instead of 10, appear to confirm this

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2010.0 ± 8.1	-0.21 ± 0.07	-0.008 ± 0.003	80.885 ± 0.260
ES	318.3 ± 2.9	0.34 ± 0.07	0.033 ± 0.007	29.449 ± 0.068
NC	2148.2 ± 8.9	-0.45 ± 0.08	-0.018 ± 0.003	89.519 ± 0.264
EN	191.8 ± 4.6	0.94 ± 0.08	0.292 ± 0.025	46.180 ± 0.082
T_{eff} scale	-0.0006 ± 0.0043	-0.088 ± 0.062	-	0.007 ± 0.000
T_{eff} offset	-0.0024 ± 0.0159	-0.046 ± 0.032	-	0.050 ± 0.000
R scale	0.0006 ± 0.0014	0.062 ± 0.014	-	0.010 ± 0.000
β_{14} scale	-0.0001 ± 0.0041	-0.025 ± 0.085	-	0.005 ± 0.000
β_{14} offset	0.0000 ± 0.0000	0.016 ± 0.009	-	0.000 ± 0.000

Table 7.4: Verification tests for an energy-constrained fit including systematic parameters. The mean fitted value, mean bias and mean fitted error are given for each parameter. For the fitted number of CC, ES, NC and EN events the fractional bias is also given.

hypothesis.

The fractional biases seen on the fitted numbers of events from this set of fits are consistent with those obtained for the fits without systematic fit parameters. Therefore, 20 bins in R^3 were used to create PDFs for the fit to salt data presented in the next section.

7.8 Energy-Constrained Fit Results

Energy-constrained fits[‖] to the salt data set were carried out treating the 5 dominant systematic uncertainties in two different ways. The techniques described in this chapter were used to perform the first fit, in which all 5 systematics were varied as parameters of the fit. Secondly, the effects of these systematics were evaluated by repeating the fit with systematically shifted PDFs, following the procedure detailed in section 6.5. The systematic corrections to observable parameters detailed in table 6.4 were not applied to the PDFs for either of these fits.[**]

[‖]Energy-constrained fits use PDFs in T_{eff} that effectively constrain the CC and ES energy distribution to that predicted for an undistorted ^{8}B spectrum. The CC and ES fluxes are each treated as a single fit parameter.

[**]This is because this analysis was performed before the corrections in the previous chapter had been finalised.

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2005.5 ± 8.1	-0.26 ± 0.07	-0.010 ± 0.003	80.937 ± 0.262
ES	318.6 ± 2.9	0.35 ± 0.07	0.034 ± 0.007	29.452 ± 0.070
NC	2160.3 ± 9.0	-0.31 ± 0.08	-0.012 ± 0.003	89.819 ± 0.276
EN	182.6 ± 4.7	0.73 ± 0.08	0.227 ± 0.026	46.605 ± 0.088
T_{eff} scale	-0.0005 ± 0.0041	-0.067 ± 0.061	-	0.007 ± 0.000
T_{eff} offset	-0.0022 ± 0.0159	-0.043 ± 0.032	-	0.050 ± 0.000
R scale	-0.0002 ± 0.0019	-0.022 ± 0.019	-	0.010 ± 0.000
β_{14} scale	0.0002 ± 0.0039	0.025 ± 0.078	-	0.005 ± 0.000
β_{14} offset	0.0000 ± 0.0000	0.018 ± 0.009	-	0.000 ± 0.000

Table 7.5: Verification tests for an energy-constrained fit including systematic parameters. The fit procedure was identical to that used to obtain the results in table 7.4 except that the R^3 PDF was created with 20 bins instead of 10. The mean fitted value, mean bias and mean fitted error are given for each parameter. For the fitted number of CC, ES, NC and EN events the fractional bias is also given.

Method 1: Varying Observable Systematics in the Fit

The bias-corrected fitted parameters are given in table 7.6. The bias correction was applied to both the fitted numbers of events and the fitted uncertainties. The uncertainties presented are those obtained from the likelihood fit, which include both statistical and systematic contributions. The preferred value for the T_{eff} scale is 2σ greater than zero, but well within the 1σ allowed range for this parameter obtained from calibration source studies. All other systematic parameters were fitted with values consistent with zero. The sum of the fitted distributions are presented in figure 7.4 in terms of the four observable parameters.

Method 2: Systematics from Shifted PDFs

The fitted parameters from this method, which have also been corrected for bias, are given in table 7.7. The separate statistical and systematic uncertainty is included for each flux, as well as the combined uncertainty, which can be directly compared to the uncertainty given in table 7.6.

7.9 Discussion

The combined uncertainty due to statistical and dominant systematic uncertainties for the two analysis methods presented in this chapter are given in table 7.8 as percentages of the

X	$F(X)$	$\sigma(X)$
CC	2134.3	$^{+76.6}_{-77.1}$
ES	283.3	$^{+28.0}_{-27.1}$
NC+int n	2172.6	$^{+87.3}_{-85.7}$
EN	144.0	$^{+42.6}_{-41.3}$
T_{eff} scale	0.012	$^{+0.006}_{-0.006}$
T_{eff} offset	0.034	$^{+0.069}_{-0.066}$
R scale	-0.003	$^{+0.014}_{-0.014}$
β_{14} scale	0.002	$^{+0.005}_{-0.004}$
β_{14} offset	0.00002	$^{+0.0004}_{-0.0004}$

Table 7.6: Results of an energy-constrained fit to the salt data set including the dominant systematic uncertainties as parameters of the fit. $\sigma(X)$ should be considered as the combined statistical and systematic uncertainty.

X	$F(X)$	Statistical	Systematic	Combined Uncertainty
CC	2232.8	$^{+70.5}_{-69.6}$	$^{+74.9}_{-75.0}$	$^{+102.8}_{-102.3}$
ES	279.2	$^{+27.9}_{-27.0}$	$^{+2.9}_{-3.3}$	$^{+28.1}_{-27.2}$
NC+int N	2103.5	$^{+77.8}_{-76.8}$	$^{+72.3}_{-72.9}$	$^{+106.3}_{-105.9}$
EN	122.9	$^{+38.6}_{-37.8}$	$^{+13.1}_{-11.1}$	$^{+40.8}_{-39.4}$

Table 7.7: Results of an energy-constrained fit to the salt data set with systematic uncertainties obtained by repeating the fit with shifted PDFs.

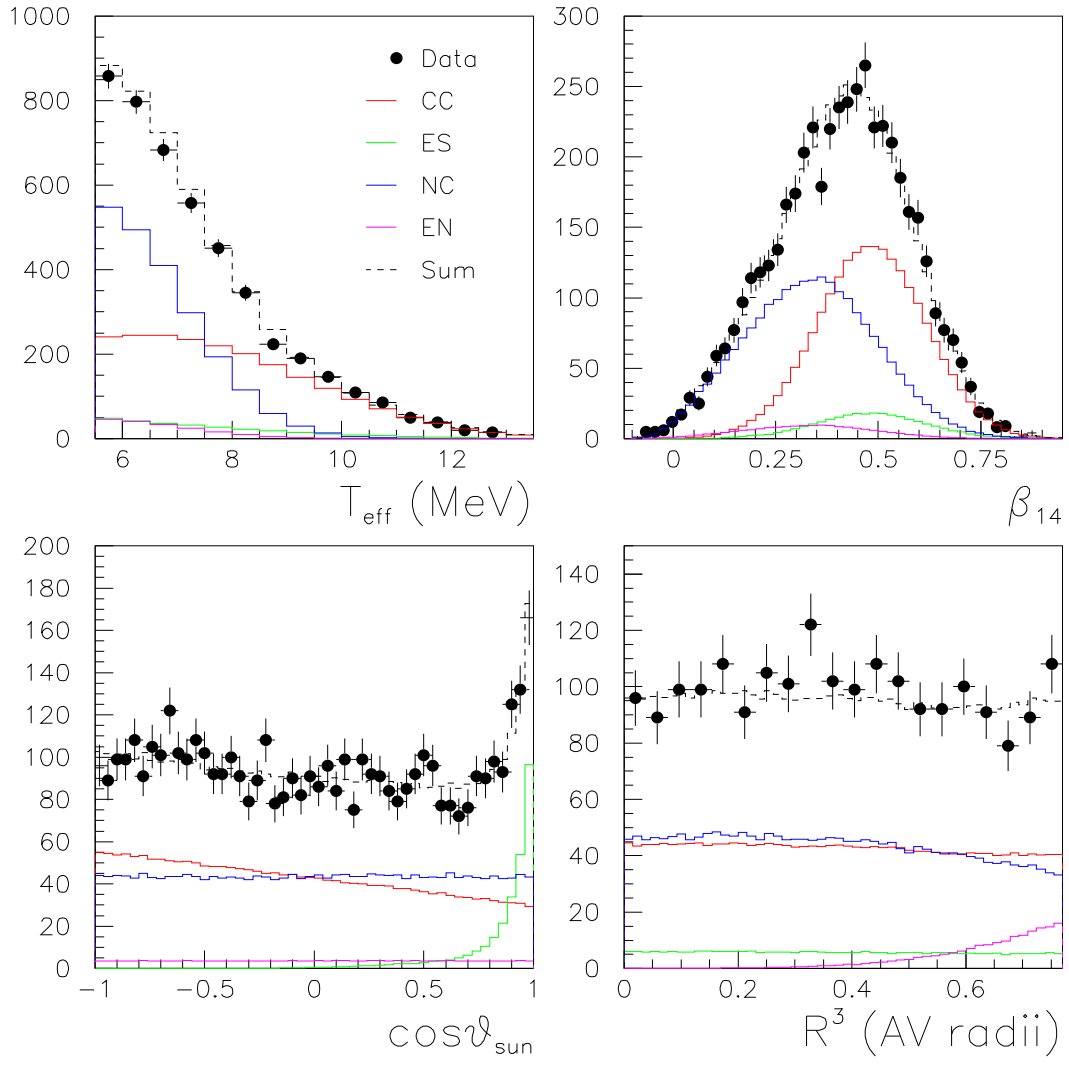


Figure 7.4: Results of the energy-constrained fit with systematic uncertainties varied in the fit, in terms of the four observable parameters used to separate events. The predicted distributions for each signal are shown scaled by the fitted number of events. The dashed line shows the sum of the fit. The data points have been shifted by the fitted systematic parameters.

	Percentage Uncertainty	
X	Variable Systematics	Shifted PDFs
CC	$^{+3.59}_{-3.61}$	$^{+4.60}_{-4.58}$
ES	$^{+9.88}_{-9.56}$	$^{+10.1}_{-9.74}$
NC+int N	$^{+4.02}_{-3.94}$	$^{+5.05}_{-5.03}$
EN	$^{+29.6}_{-28.7}$	$^{+33.2}_{-32.1}$

Table 7.8: Comparison of the total uncertainty contributions (statistical and dominant systematic uncertainties) from the two different approaches to systematic uncertainty propagation.

fitted numbers of events. For all fit parameters, the combined uncertainty obtained from the fit including variable systematic parameters is noticeably smaller. The uncertainty for CC events is reduced by 25%, and the for NC events the uncertainty is reduced by 20%. This indicates that incorrect treatment of the correlations between the systematic uncertainties on observable parameters, results in a overly conservative estimate of the uncertainties in the fitted numbers of events, and consequently in the CC/NC ratio that is used to test against different MSW oscillation scenarios.

7.10 Summary

In this chapter, a method to include systematic parameters as variables in the likelihood fit has been described. A number of difficulties in this new approach were overcome and the tools developed to perform this type of fit have been proved to work well for energy-constrained likelihood fits. There are additional complications associated with applying this method to an energy-unconstrained fit, which have not been addressed for the purposes of this thesis.

The new method of systematic uncertainty propagation was compared to the method of shifting PDFs adopted in the previous chapter, and was found to reduce the combined uncertainty on the number of events fitted in an energy constrained fit to the salt data set.

Chapter 8

The Electron Neutrino Energy Spectrum

The energy spectrum of neutrinos arriving at the SNO detector could be distorted under various oscillation scenarios. The charged-current spectrum presented in chapter 6 can be used to test for these distortions. This measurement was presented in terms of T_{eff}, the observed energy but the fundamental parameter of interest is E_ν, the incoming neutrino energy. As the energy of simulated neutrino interactions is dictated by the initial neutrino energy spectrum, E_ν information is contained in the likelihood function. A simple change in variables can be made to perform the fit in terms of this parameter. This chapter explores the measurement of the "neutrino energy" spectrum.

8.1 Maximum Likelihood Fit Parameters

Matter effects would serve to suppress the flux of electron-type neutrinos from the Sun in an energy-dependent manner. The most direct evidence for this would be a change in the electron neutrino spectrum with respect to the undistorted predicted spectrum. To probe this, the signal extraction procedure was adapted so that the parameters varied in the fit were the total flux of ^{8}B neutrinos, $F_{\rm B}$ defined relative to a predicted flux* of $5.15 \times 10^6\,{\rm cm}^{-2}\,{\rm s}^{-1}$, and the fraction of neutrinos that are electron-type in each neutrino energy bin.

These parameters were used to derive the numbers of CC, ES and NC events, as given in equations 8.1 to 8.3.

$$N_i(CC) = F_{ei} \times E_i(CC) \tag{8.1}$$

*This is the prediction for the ^{8}B flux used in the Monte Carlo simulations. *hep* neutrinos also make a small contribution to $F_{\rm B}$ but this value is small (<0.2% of the ^{8}B predicted flux).

$$N_i(ES) = (F_{ei} + \epsilon_{\rm ES}(1 - F_{ei})) \times E_i(ES) \tag{8.2}$$

$$N(NC) = F_{\rm B} \times E(NC) + N_{\rm INB} \tag{8.3}$$

In these equations F_{ei} is the fraction of electron-type neutrinos detected in bin i, which is a variable of the fit, and the parameters $E_i(j)$ give the number of predicted events in that neutrino energy bin for signal j. For the CC and ES signals these values were obtained from Monte Carlo simulations and included the corrections detailed in appendix I, with the exception of the model correction factor, $R_{\rm model}$. The calculation of $E(NC)$ is given in equation I.5 though, again, the model correction factor was omitted. $\epsilon_{\rm ES}$ is the relative sensitivity of the ES signal to ν_μ and ν_τ neutrinos with respect to ν_e. $N_{\rm INB}$ is the internal neutron background due to photo-disintegration which is indistinguishable from the NC signal. The internal neutron background was treated as a fixed parameter in the fit.

$\epsilon_{\rm ES}$ has a slight energy dependence as the cross sections for ν_e and $\nu_{\mu\tau}$ scattering are dependent on the kinetic energy of the scattered electron. In theory, this would also change the shape of the ES T_{eff} distribution for the different neutrino flavours, but the combination of the kinematics of elastic scattered events and the detector energy resolution make this a negligible effect. The simulated events used for creating PDFs were all generated as electron-type neutrinos. For this analysis the energy dependence of $\epsilon_{\rm ES}$ was assumed to be negligible and a value of 0.1559, obtained from integration of the differential cross-sections above 5.5 MeV [107] was used.

An implicit assumption of this approach is that the underlying neutrino energy spectrum used to generate the simulated events which make up the PDFs, is correct.[†] The undistorted ^{8}B spectrum calculated in Ref. [29] and shown in figure 8.1 was used for all Monte Carlo simulations of neutrino events.

8.2 Application in MXF

The parameter conversions described in the previous section were applied in the MXF code using values of $N_i(j)_{MC}$, required for the calculation of $E_i(j)$, obtained from the normalisation of PDFs. Changes in normalisation due to systematic fit parameters were naturally included in this way. $E(NC)$ was also corrected for relative changes in normalisation when systematic parameters were included in the fit. The log likelihood function

[†] There are theoretical uncertainties associated with the model that have been omitted for the purposes of this analysis.

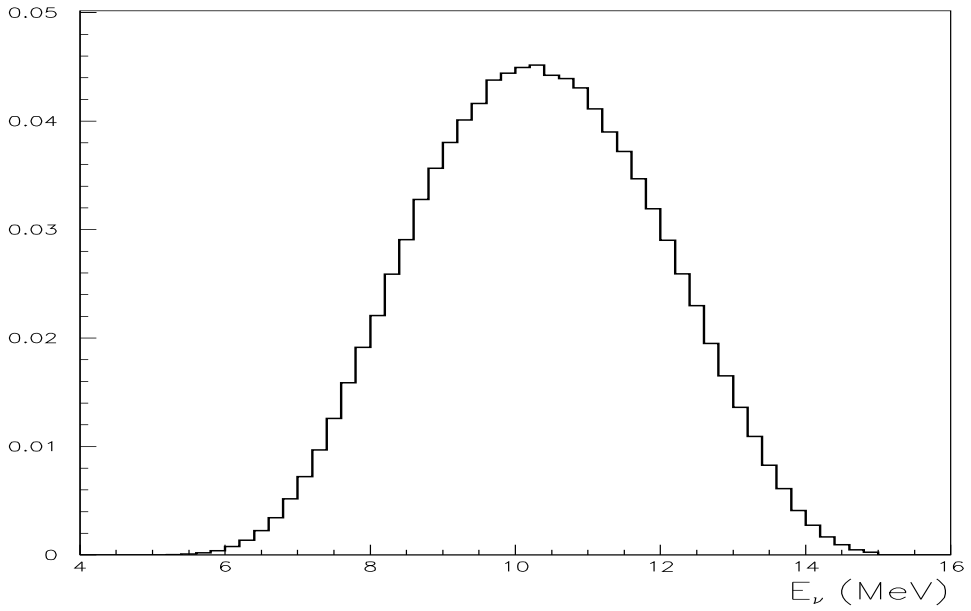


Figure 8.1: The energy spectrum of ^{8}B neutrinos assumed for Monte Carlo event simulation.

was determined from the calculated values of $N_i(j)$ as given in 8.4.

$$\mathbb{L} = \sum_{d=1}^{N_{\rm data}} log\left(\sum_{i=1}^{N_{\rm s}} N_i P(i/T_{eff}, \beta_{14}, \cos\theta_\odot, R^3)\right) - \left(\sum_{i=1}^{N_{\rm s}} N_i\right) \tag{8.4}$$

where $N_{\rm data}$ is the number of events in the data set and $N_{\rm s}$ is the number of signals. The number of events in each CC and ES neutrino energy bin, and the number of NC, EN, INB and internal γ events were all treated as separate signals in this context. N_i is the number of events of signal type i, calculated as prescribed by equations 8.1 to 8.3 for the neutrino signals.

The probabilities specified in equation 8.4 were obtained from PDFs created separately for each neutrino energy bin for CC and ES signals, whilst a single PDF was created for the NC signal and each of the backgrounds. The PDFs in observable energy, T_{eff}, are shown for the CC and ES signals in figure 8.2. A total of 7 bins in neutrino energy were selected with widths as specified in table 8.1.

The factorisation of PDFs (specified in equation 5.21), used in the previous chapter, was also adopted for the analysis presented in this chapter. Although this factorisation does not fully account for all the correlations between observable parameters, the magnitude of the bias introduced is known to be small and systematics can be more easily

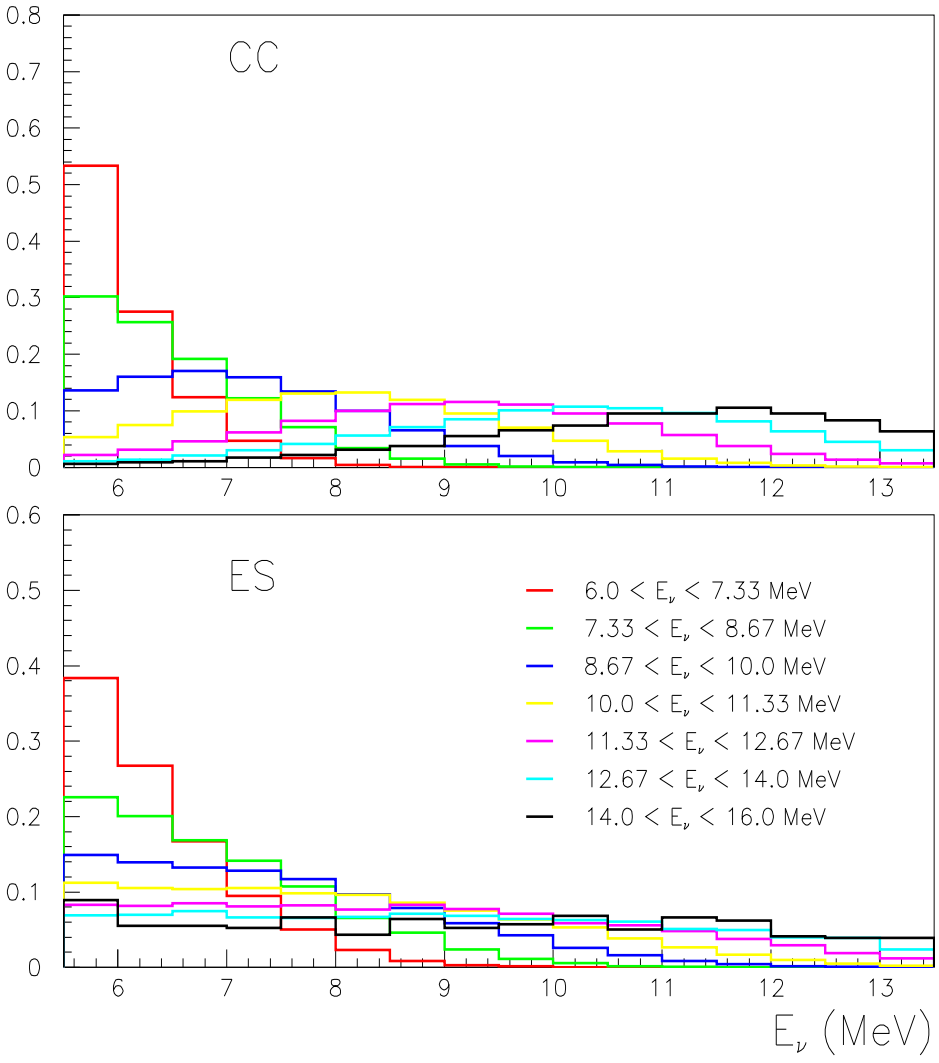


Figure 8.2: PDFs in observed kinetic energy, T_{eff} for different neutrino energy bins for the CC and ES signals.

Bin Number	Range in E_ν (MeV)
1	4.0–7.33
2	7.33–8.67
3	8.67–10.0
4	10.0–11.33
5	11.33–12.67
6	12.67–14.0
7	14.0–16.0

Table 8.1: Bins sizes chosen in neutrino energy.

propagated when fitting with lower dimensionality PDFs. Therefore, the correction to account for the biases, given in table 7.2, was applied to the predicted number of events for each signal. As there is only weak evidence for a differential bias in T_{eff} and the expected values are for neutrino energy bins, rather than T_{eff} bins, the same bias correction was applied to each expected flux value, $E_i(j)$.

8.2.1 Total Electron Fraction

To obtain F_e, the overall ratio of the electron-type neutrino flux to the model prediction, from the fitted F_{ei} values, correlations between these parameters must be considered. Instead of performing this non-trivial calculation on the parameters returned by the likelihood fit, the variables of the fit were changed to include the value F_e directly.

For a fit in f neutrino energy bins, the fit parameter F_{ef} was replaced by F_e.[‡] The value, F_{ef}, required to calculate the number of events for each signal in the final bin was calculated in the MXF code, as detailed below.

The number of electron-type neutrino events in the ith bin, ϕ_{ei}, is given by:

$$\phi_{ei} = Q_{\rm SM}(i) F_{ei}, \tag{8.5}$$

where $Q_{\rm SM}(i)$ is the predicted number of electron-type events in that bin:

$$Q_{\rm SM}(i) = E_i(CC) + E_i(ES). \tag{8.6}$$

Similarly, the total number of electron-type neutrino events is:

$$\phi_e = Q_{\rm SM} F_e \tag{8.7}$$

where $Q_{\rm SM}$ is the total predicted number of electron-type neutrino events, obtained from the summation of $Q_{\rm SM}(i)$ over f bins. $E_i(CC)$ and $E_i(ES)$ are the predicted numbers of

[‡]This was an optional change in variables. The fit was carried out separately to obtain the full neutrino energy spectrum

CC and ES events, respectively, in bin i. The flux, ϕ_{ef}, in bin f is given by:

$$\phi_{ef} = \phi_e - \sum_{i=1}^{f-1} \phi_{ei}. \tag{8.8}$$

Thus, the value of F_{ef} is obtained from equation 8.9:

$$F_{ef} = \frac{1}{Q_{\rm SM}(f)} \left[Q_{\rm SM} F_e - \sum_{i=1}^{f-1} Q_{\rm SM}(i) F_{ei} \right] \tag{8.9}$$

8.3 Deconvolution

Figure 8.2 shows a high degree of overlap between the T_{eff} distributions for separate E_ν bins. Therefore, the fitted parameters, F_{ei}, will be highly correlated. This overlap is due to both the limited resolution of the SNO detector and the kinematics of the neutrino interactions which both serve to "smear" the observed spectrum. Deconvolving these effects from the underlying neutrino energy spectrum provides a result that can be directly compared with other experiments. However, deconvolution can result in unreal scenarios which give good fits to the data, such as alternate bins with large and small amplitude. In such situations, when a range of parameters give equally good fits to the data, the uncertainties on the fit parameters are, naturally, large. Such effects are seen in the electron neutrino spectrum obtained for the salt data set from this approach. However, the total fluxes are unaffected by these correlations.

8.4 The Neutrino Energy Spectrum

The neutrino energy spectrum, obtained for the salt data set is shown in figure 8.3. For the purposes of this analysis, no systematic effects on observable parameters were considered and, therefore, only statistical uncertainties are shown on this figure. The effects of deconvolution are clearly visible in the fluctuation of points and the large statistical uncertainties although the fitted CC spectrum, shown in figure 8.4 looks very reasonable. This spectrum was obtained by weighting the CC PDFs (shown in figure 8.2) for each neutrino energy bin by the predicted flux in each bin weighted by the fitted F_{ei} value and is compared to the undistorted ^{8}B spectrum shape, also normalised to the total number of fitted CC events. The extracted spectrum is lower than expected at low energies, following the same trend as the spectrum presented in chapter 6 but the uncertainties at low energy (which have been omitted in this figure) mean that this dip is not very significant.

As expected, the deconvolution does not appear to affect the total neutrino flux ratio, which was fitted as:

$$F_{\rm B} = 1.072 \pm 0.04 \text{ (stat.)}$$

This was converted to a total flux of solar neutrinos of

$$\Phi_x = 5.52 \pm 0.21\,(\text{stat.}) \times 10^6\,\text{cm}^{-2}\,\text{s}^{-1}.$$

Using the method detailed in section 8.2.1, the fitted electron-type flux fraction, which is shown as a band on figure 8.3, was:

$$F_e = 0.331 \pm 0.012\,(\text{stat.}).$$

This relates to an electron-type neutrino flux of

$$\Phi_e = 1.705 \pm 0.062\,(\text{stat.}) \times 10^6\,\text{cm}^{-2}\,\text{s}^{-1}.$$

As expected the uncertainty on the electron fraction in individual neutrino energy bins is large but the uncertainty on the overall electron-type flux is small. The fitted flux ratios were converted into numbers of events for the three types of neutrino interaction and are given in the first column (Fit A) of table 8.2. For comparison, the fit values obtained for a direct energy-unconstrained fit similar to that presented in section 6.7 are given in the second column (Fit B).[§]

The statistical uncertainties obtained from the likelihood fit were propagated through equations 8.1 to 8.3 to obtain the uncertainties on the parameters from Fit A.[¶] Fit B was carried out with 3D factorised PDFs that account for correlations between the observable distributions, whilst Fit A was corrected for biases due to PDF factorisation using the corrections given in table 7.2.

The results obtained from the two fit approaches are in very good agreement. The uncertainty on the number of ES events from Fit A is significantly lower than that obtained for Fit B. This is due to the theoretical constraints applied to the relationship between the three signal types.

The direct fit to the number of each event type results in a 3.7% error on the fitted number of CC events which can be considered as a measure of the integral electron-type flux. The fractional error obtained on F_e is $\pm 3.6\%$. The fractional error on the fitted number of NC events is 3.9% which is used as a measure of the total neutrino flux, whilst the fractional error on F_{B} is 3.7%. Therefore, the fit method presented in this chapter, achieves a modest reduction in the statistical uncertainties on the parameters of interest.

8.5 Hypothesis Testing

Due to the problems of deconvolution, it is hard to use electron neutrino spectral information without making prior assumptions. However, the conversion from the number of

[§]In this case, no corrections were applied to the PDFs for observable systematic effects.

[¶]With the exception of the EN uncertainty, which is a direct output of the likelihood fit.

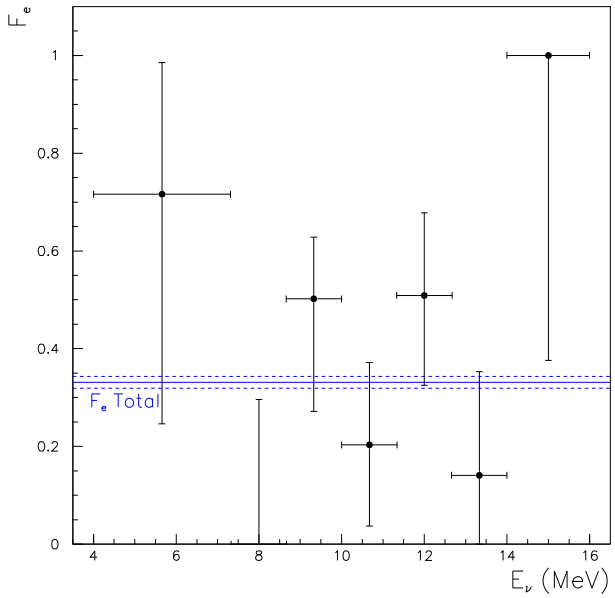


Figure 8.3: Fitted shape of the electron neutrino energy spectrum with statistical errors only. The band indicates the overall reduction factor of electron-type neutrinos with respect to a predicted flux of $5.15 \times 10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

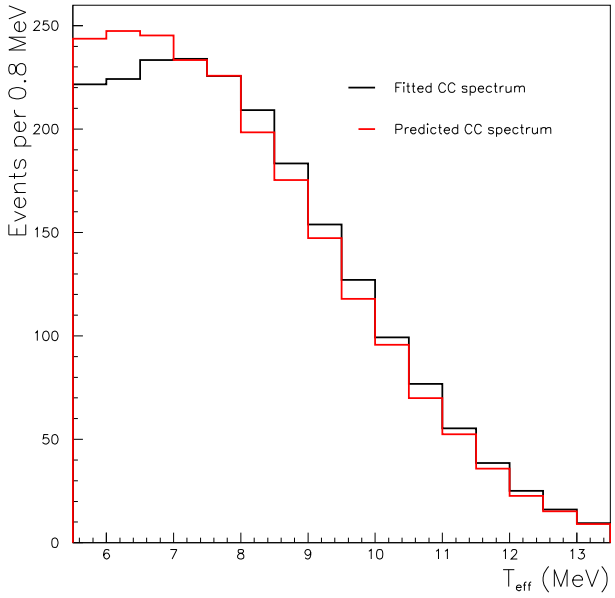


Figure 8.4: The CC spectrum obtained using the fitted F_{ei} values shown in figure 8.3 is shown in black. The predicted CC spectrum for undistorted ^{8}B events is shown in red and has been normalised by the same total number of CC events.

Signal	Fit A	Fit B
CC	2167.3±78.6	$2146.9^{+79.4}_{-78.5}$
ES	269.0±8.2	$282.3^{+28.3}_{-27.3}$
NC+internal n	2172.7±87.6	$2181.7^{+85.3}_{-84.4}$
EN	$126.6^{+38.8}_{-37.8}$	$127.6^{+42.9}_{-41.8}$
Internal γs	8.5	8.5
Total	4744	4747

Table 8.2: Fitted numbers of events from a fit in neutrino energy (Fit A) compared to results obtained from a direct fit to the number of each event type (Fit B). Statistical errors only are given on the Fit B parameters.

measured events for each signal to true neutrino flux parameters is a useful technique. This section discusses how such a technique can be applied to a test of different model spectra.

Fits which use information from T_{eff} PDFs to separate events naturally constrain the neutrino energy distribution. All the PDFs discussed so far were created from simulations of events due to electron-type neutrinos with an undistorted ^{8}B energy spectrum. For the analysis described in this section, PDFs with the energy distribution predicted by MSW effects were created for various mixing parameters using the tools described in section 5.5.2. These MSW-weighted PDFs not only provide the predicted energy distribution for each type of signal, their normalisation can also be used to predict the relative numbers of events once the corrections described in appendix I and section 8.2 have been applied.

The salt data set was tested against a range of MSW mixing scenarios by constraining both the energy distribution and the flux of electron-type neutrinos in the likelihood fit. The fits were carried out in a single neutrino energy bin, with the fraction of electron-type neutrinos constrained to be that predicted by the model. The relative maximum likelihood values for the different models were then used to indicate the relative probability for each hypothesis tested using a likelihood ratio.

The flux conversions used for this analysis were:

$$N(CC) = F_{\rm B} \times E(CC), \tag{8.10}$$

$$N(ES) = F_{\rm B} \times E(ES), \tag{8.11}$$

$$N(NC) = F_{\rm B} \times \delta \times E(NC) + N_{\rm INB}, \tag{8.12}$$

where the expected numbers of events, E, was still obtained from PDF normalisation for the CC and ES signals, but these normalisations included a weighting to account for the survival probabilities of the tested mixing scenario. The scale factor for the total flux of ^{8}B neutrinos, $F_{\rm B}$, was allowed to vary in the fit. Since this factor is now involved in the

flux conversion for each signal type, uncertainties in one flux conversion are correlated with uncertainties in the other fluxes. For this reason, the factor δ in equation 8.12 was introduced to account for systematic uncertainty in the neutron capture efficiency. δ was allowed to vary as a parameter of the fit, but was constrained about unity by the 1σ uncertainty estimates on the parameter $\epsilon_{\rm n}$, given in appendix I.

An error in $N_{\rm INB}$ would also affect all the fluxes in this scenario so the value of $N_{\rm INB}$ was allowed to vary in the fit, constrained by 1σ uncertainty estimates obtained from calibration studies (see equation 6.15).

Systematic changes in the observable distributions will also affect the predicted number of events for each signal so ideally systematic parameters should be allowed to vary in the fit. However, initial attempts to incorporate systematic parameters in the fits described here were not satisfactory. The problem is thought to be due to local minima in likelihood space caused by the correlations between the fit parameters which determine the number of events of each signal type. (*eg.* In this format, the number of NC events used in the likelihood calculation is dependent on all the systematic parameters and also the three variables: $F_{\rm B}$, δ, $N_{\rm INB}$.) This problem requires further investigation but for the purposes of this thesis, the method of scanning MSW space is presented without consideration of systematic effects.

8.6 A Scan of MSW Parameter Space

The region of Δm^2 – $\tan^2\theta$ parameter space relating to the LMA solution, the preferred mixing parameters for solar neutrino oscillations, is shown in figure 1.4. The fit procedure described in the previous section was applied to the salt data set for a scan of this MSW parameter space. 32 values of Δm^2, equally separated in $\log_{10}\Delta m^2$ were sampled in the region 10^{-5}–$10^{-3.45}$, and 15 values of $\tan^2\theta$, equally spaced in the range 0.2–0.9, were sampled.

The parameters $\Delta m^2 = 2.0\times10^{-5}\,{\rm eV}^2$ and $\tan^2\theta = 0.55$ were found to give the highest maximum likelihood value. The likelihood ratio, λ, as described in section 5.3, was calculated for each fit with respect to this best fit point. The values falling within 1σ ($\lambda = 0.5$, black), 2σ ($\lambda = 2.0$, red), 3σ ($\lambda = 2.0$, green) and 4σ ($\lambda = 2.0$, blue), are shown in figure 8.5. The small dots in this figure indicate points sampled which fall outside the 4σ confidence level.

This figure can be compared to a contour plot of allowed parameter values obtained from a χ^2 fit to all SNO data (salt and D_2O phase), which is shown in figure 8.6. Only the LMA region is shown in this figure, although some other mixing solutions are only ruled out when information from the radiochemical experiments is included in the fit.

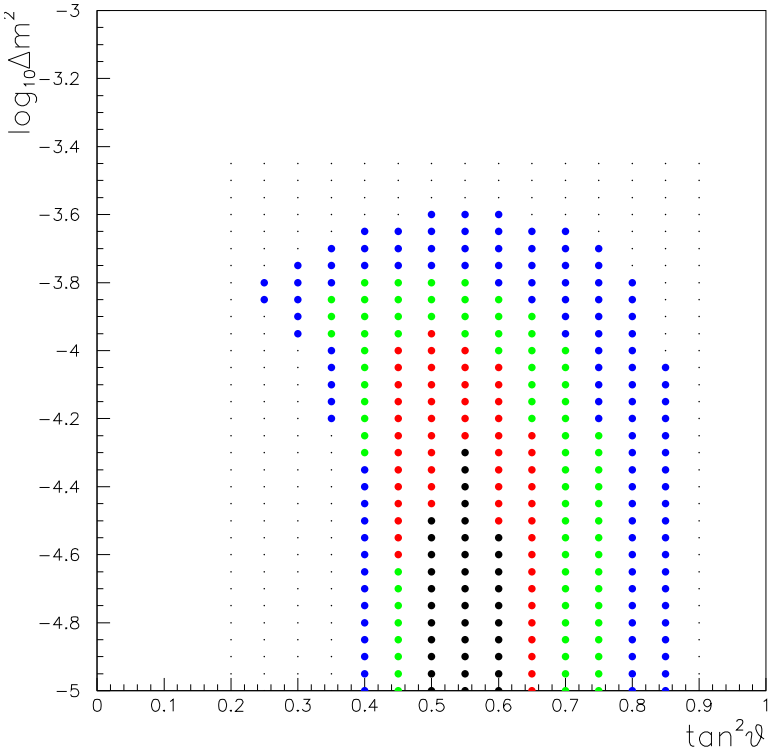


Figure 8.5: The likelihood ratio for a scan of LMA MSW parameter space. The values falling within 1σ ($\lambda = 0.5$, black), 2σ ($\lambda = 2.0$, red), 3σ ($\lambda = 2.0$, green) and 4σ ($\lambda = 2.0$, blue) of the best fit point are plotted as full circles. The small dots indicate all the points sampled which fall outside the 4σ confidence level.

Note that this figure, is shown on a logarithmic scale and includes information from both CC spectrum and day-night analyses. Systematic uncertainties have also been taken into account in this contour plot, but not in figure 8.5.

The χ^2 contours for the LMA region obtained from the likelihood ratio test differ from those in figure 8.6 at lower values of Δm^2. This is where the day-night asymmetry, which was not considered in the likelihood ratio test, has the strongest affect.

The contours obtained from the MXF approach are slightly narrower in $\tan^2\theta$, which may be due to the omission of systematic uncertainties. However, extra information is included in this likelihood approach which uses PDFs in β_{14}, R^3 and $\cos\theta_\odot$ as well as the T_{eff} information used in the χ^2 minimisation. It is anticipated that varying systematics as parameters in the fit in the likelihood test would help to reduce the allowed parameter space. Further investigation is required for this procedure and also to combine the salt and D_2O data sets in this framework.

8.7 Summary

In this chapter techniques to measure the true "neutrino energy" spectrum have been presented. The correlations between fitted electron-type neutrino fluxes in separate E_ν bins are large, but the flux uncertainties obtained from an energy-unconstrained fit in terms of this parameter are comparable to, if not marginally smaller than, uncertainties obtained on the fitted CC and NC fluxes presented in chapter 6.

Tools were also developed to test the salt data set against the predictions of MSW models for a range of mixing parameters in the maximum likelihood framework. This method included information from all observable parameters, unlike χ^2 analyses, which only used T_{eff} information. Future work to incorporate systematic parameters in this fit, using the procedure described in chapter 7, will show whether this method can be used to improve on the limits for allowed parameter space. Combining this analysis with other SNO data such as day-night asymmetries may also help to improve the limits SNO can place on the mixing parameters Δm_{12} and θ_{12}.

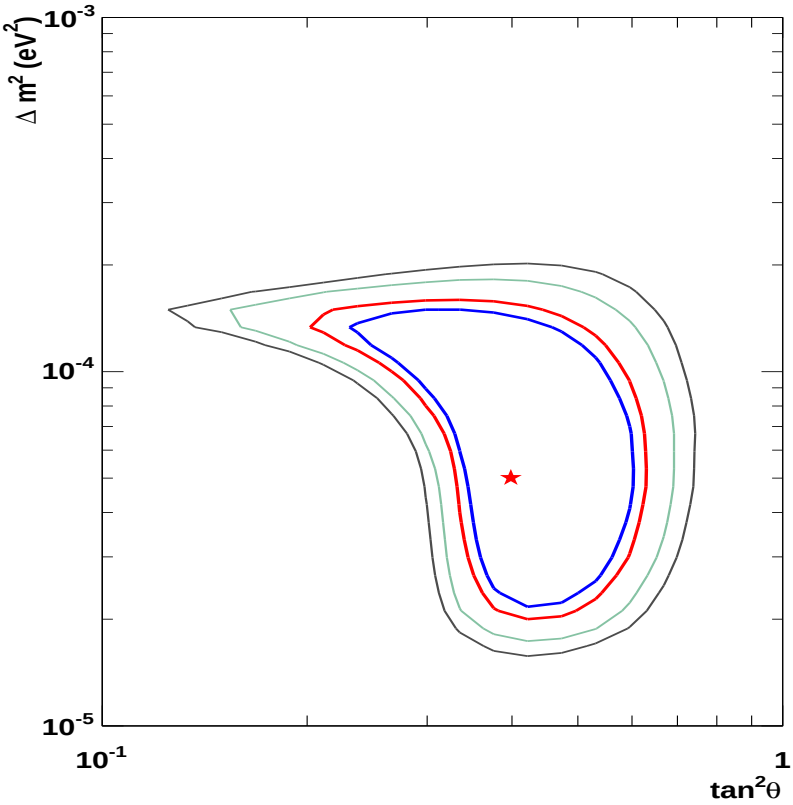


Figure 8.6: Allowed MSW parameter space obtained from a χ^2 fit to SNO D_2O day and night spectra, salt day and night spectra, ES and NC day and night fluxes. The contours represent χ^2 confidence levels relating to 90% (blue), 95% (red), 99% (green) and 99.73% (black) probability. This is a preliminary plot kindly generated by S. Peeters from information obtained from Ref. [106].

Chapter 9

Conclusions and Discussion

In this thesis, a number of analyses of the SNO salt data set have been presented. Maximum likelihood signal extraction was developed to perform these analyses with a high degree of flexibility and functionality as discussed in chapter 5. In chapters 3 and 4 studies performed to characterise elements of the energy and optical response of the SNO detector were presented. Rigorous calibrations like these have allowed stringent limits to be placed on the magnitude of systematic uncertainties in the observable distributions used to separate the different neutrino interactions detected in SNO.

In chapter 6, a measurement of the charged-current energy spectrum was described, including a full treatment of systematic uncertainties. The spectrum is consistent with the undistorted ^{8}B shape, showing no evidence of a modification which could result from matter effects in ν oscillations. However, the statistical uncertainty alone in this measurement is larger than the expected distortion for the LMA MSW scenario. No evidence for a day-night asymmetry was found either as the separate day and night spectra and fluxes measured were in good agreement.

As the full salt data set was analysed here, the statistical uncertainties in this measurement can only be reduced by combining data from other phases of the SNO experiment. More than 300 live-days of data are available from phase I of the experiment, although careful consideration of correlations between the systematic uncertainties from the different phases will be required. The NCD phase, which is now commencing, will provide further data that may also help to constrain the spectrum measurement as the majority of neutron events will be detected in a separate data stream in this phase. More pronounced distortions to the energy spectrum are predicted at lower energies for the LMA mixing scenario so ideally the spectrum measurement should be extended to a lower energy threshold. However, the magnitude of systematic uncertainties in the lowest energy bin of the presented CC spectrum indicate that little useful information will be obtained from SNO data in this systematically dominated region.

Total fluxes were also obtained from this energy-unconstrained fit. The total flux of active neutrinos, measured from NC events was found to be:

$$\Phi(\mathrm{NC}) = (5.23 \pm 0.23(\mathrm{stat.}) \pm 0.27(\mathrm{sys.}) \pm 0.06(\mathrm{theory})) \times 10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}. \quad (9.1)$$

This was obtained in a manner which was independent of assumptions of the neutrino-mixing mechanism and was found to be in good agreement with the predicted solar neutrino flux [21]. The ratio of fluxes due to CC and NC interactions was calculated to be:

$$\frac{\Phi(\mathrm{CC})}{\Phi(\mathrm{NC})} = 0.310 \pm 0.021(\mathrm{stat.}) \pm 0.024(\mathrm{sys.}), \quad (9.2)$$

which is in good agreement with that previously published for a sub-set of the salt data [31], but with reduced statistical uncertainty. This improved ratio, along with the lack of evidence for spectral distortions or day-night asymmetry effects, was used to constrain the allowed parameter space for oscillations between the ν_1 and ν_2 neutrino mass states.

In chapter 7, a different approach to systematic uncertainties was explored in the context of an energy-constrained fit. Inclusion of systematic parameters in the likelihood fit formally accounted for correlations between fit parameters and reduced the combined error on the fitted numbers of CC and NC events significantly (20–25%). Application of this method to the CC spectrum measurement (an energy-unconstrained fit) could further reduce uncertainties, although there are complications associated with the use of binned PDFs which must be overcome.

In chapter 8 measurement of the spectrum in terms of the initial neutrino energy was explored. The neutrino energy spectrum obtained was hard to interpret because the uncertainties in this measurement were strongly correlated due to the effects of deconvolution from the detector response. However, the uncertainties on the total fitted fluxes were comparable and, in the case of ES events, smaller than those obtained from a direct fit to the numbers of CC, ES and NC events because of the theoretical constraints between the different interaction rates which were incorporated in the likelihood function. This technique was also used to perform a scan in MSW parameters testing the data against a series of model predictions. The results of this scan were in good agreement with allowed contours in parameter space obtained from global analyses of SNO data and the method could be improved by incorporating systematic parameters using the procedures described in chapter 7. Merging the likelihood maps obtained from such hypothesis tests with data from other experiments, and also SNO day-night asymmetry measurements may serve to further constrain the mixing parameters.

9.1 Future Solar Neutrino Oscillation Measurements

The SNO experiment has confirmed the neutrino oscillation hypothesis through measurement of both the electron neutrino flux and the total flux of active neutrinos produced in the Sun, thus proving that neutrinos have finite mass. Data from the salt phase of the experiment, in conjunction with results from the KamLAND reactor experiment has dramatically reduced the allowed parameter space for 1-2 mixing to a single region called the LMA solution.

Future data from solar experiments such as SNO will help improve the precision measurements of the oscillation parameters Δm^2_{12} and, in particular, θ_{12}. This will allow more specific tests of neutrino oscillation theories, as well as placing stringent limits on the magnitude of sub-dominant effects such as mixing to sterile neutrino species, RSFP and neutrino decay. Combination of the measurements of solar ν_e with the KamLAND measurements of $\bar{\nu}_e$ also provides restrictions on possible CPT violation.

To date, no positive signature of matter effects has been directly measured in a single experiment. The predicted magnitude of spectral distortion and day-night asymmetry effects is small for the confirmed LMA oscillation scenario and a positive measurement of these signatures will be difficult. However, improved measurement of the energy spectrum by SNO will test other theories that also predict spectral distortions. For example, under certain scenarios, the presence of sterile neutrinos could decrease the flux of electron-type neutrinos at low energies [108].

Increased accuracy in solar neutrino measurements will also lead to an improvement in the knowledge of the other parameters in the full MNS mixing matrix (equation 1.23). As θ_{13} is small, its effect on solar oscillations is mild but, as the accuracy of data is increased, even these small effects have to be taken into account. The upper bound on θ_{13} is dominated by the CHOOZ reactor experiment [35] but measurements by solar neutrino experiments and KamLAND actually provide better constraints on this parameter for the lowest allowed values of Δm^2_{23}, as the effect of θ_{13} is to decrease the energy dependence of the solar survival probability [109]. The precision of the day-night measurements from solar experiments is currently insufficient to probe the value of θ_{13}, but future large water Čerenkov experiments may provide useful information [110].

As the solar oscillation phenomenon becomes more clearly understood, further information on the neutrino processes in the Sun can be obtained. The SNO measurements are in good agreement with the solar model predictions, but only measure the highest energy ^{8}B and *hep* neutrinos. The GNO radiochemical experiment is sensitive to ^{7}Be and pp neutrinos and is still collecting data that may improve the integral neutrino flux measurements in this energy region. The BOREXINO experiment [111] aims to commence operation in the near future, with the main objective of measuring ^{7}Be neutrinos. There

are also plans to reduce background levels in the KamLAND experiment for sensitivity to ^{7}Be neutrinos. These liquid scintillator experiments require very low levels of ^{210}Pb, ^{85}Kr and ^{40}K as well as U and Th chain elements in order to measure neutrinos with energies $\leq 1\,\mathrm{MeV}$.

9.2 The Bigger Picture

Neutrinos are one of the fundamental constituents of matter that are described by the Standard Model of Particle Physics. Whilst this model has been incredibly successful in explaining and predicting many of the properties and interactions of the fundamental particles, new theories are being developed to improve upon it and give a fuller understanding of the basic components of the Universe. The aim of these new theories is to unify all four fundamental forces, including gravity, and explain the input parameters of the existing model. The predictions of these new theories can be tested against precision measurements of a number of values, including the neutrino masses and mixing parameters. However, rigorous tests will require a high degree of accuracy in the experimental measurements.

Significant progress has been made in constraining the parameters of the MNS matrix that govern mixing between the three active neutrino flavours. In addition to the constraints on Δm_{12} and θ_{12} described in this thesis, the parameters governing atmospheric neutrino oscillations, $|\Delta m_{23}|$ and θ_{23}, are well constrained by data from SuperK and have been confirmed by the long-baseline accelerator experiment K2K [112]. The MINOS [113] experiment will commence data taking from a long-baseline neutrino beam shortly and should further constrain the allowed mixing parameters by measuring a positive oscillation signature.* This long-baseline makes MINOS sensitive to matter effects, which could provide information on the sign of Δm_{23}.

θ_{13} is the last unknown angle in the MNS matrix. At present, only upper bounds exist for this parameter but future experiments aim to measure this angle more accurately. Proposed reactor experiments will probe the value of θ_{13} through the disappearance of $\bar{\nu}_e$ at short distances (1–2 km) from nuclear reactors [114], whilst new accelerator experiments will search for ν_e appearance from a ν_μ beam. The latter will operate over long baselines (hundreds of km) in an off-axis configuration that provides a beam of ν_μ in a narrow energy range. These two types of experiment are complementary as only the accelerator experiments are sensitive to matter effects.

The measurements of θ_{13} are strongly motivated by the possibility of CP violation in the lepton sector, which may help to explain the matter-antimatter asymmetry in the

*MINOS looks for a dip in the predicted energy spectrum of the ν_μ neutrino beam.

Universe. CP violation has been observed for quarks but the magnitude of this effect is insufficient to explain this phenomenon [115]. The CP phase, δ, in the MNS matrix is only accessible if $\theta_{13} \geq 0$. Measurement of this phase would involve collecting data with long-baseline accelerator experiments from both ν_μ and $\bar{\nu}_\mu$ beams.

The MNS matrix fails to explain all experimental evidence for neutrino oscillation because the positive oscillation signature of the LSND experiment cannot be incorporated in the three neutrino oscillation scenario. If the MiniBooNE experiment confirms this result, the model of neutrinos would need to be extended to explain a fourth type of flavour change, either through other neutrino phenomena or the presence of sterile neutrino species.

The nature of neutrinos and the mechanism for production of neutrino masses is also an unresolved issue which has significant implications for future models of particle physics. One of the most promising ways to probe the nature of neutrinos is through searches for neutrinoless double beta decay, which can only occur for Majorana-type neutrinos. The current oscillation results offer a discovery potential for the next generation of experiments that search for this phenomena. Low radioactive background levels are vital in these experiments which will be built on a large scale with several hundred kilograms of material containing the candidate isotopes for this decay. Neutrinoless double beta decay measurements can also place limits on the mass of Majorana-type neutrinos and may help to determine the hierarchy of neutrino masses.

Neutrinos are the most weakly interacting particles and huge experimental advances have been required to measure their properties. However, these measurements have been extremely revealing, impacting not only on particle physics, but also astrophysics and cosmology. The measurement of the remaining unknown parameters will be even more challenging but hopefully will be equally enlightening.

Appendix A

Radioactive Decay Schemes

The decay chains of ^{232}Th and ^{238}U are displayed on the following pages.

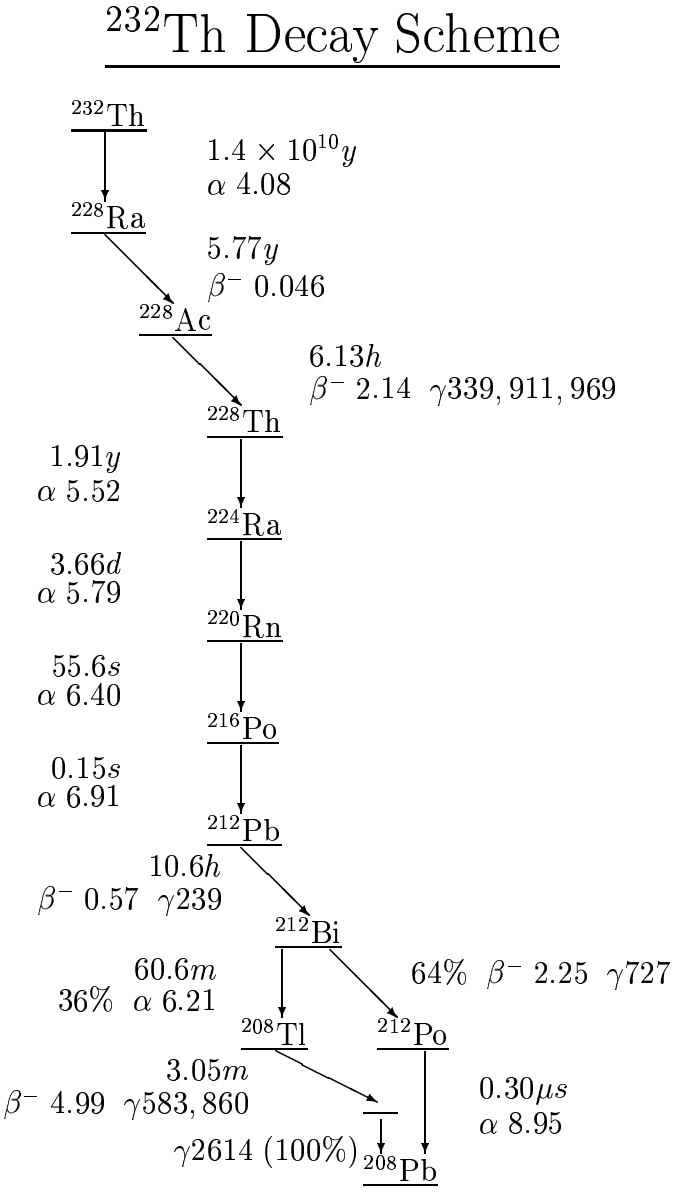


Figure A.1: The decay chain of ^{232}Th. All half-lives are shown, with the Q-values of beta and alpha decays in MeV, and gamma rays in keV. Taken from [116].

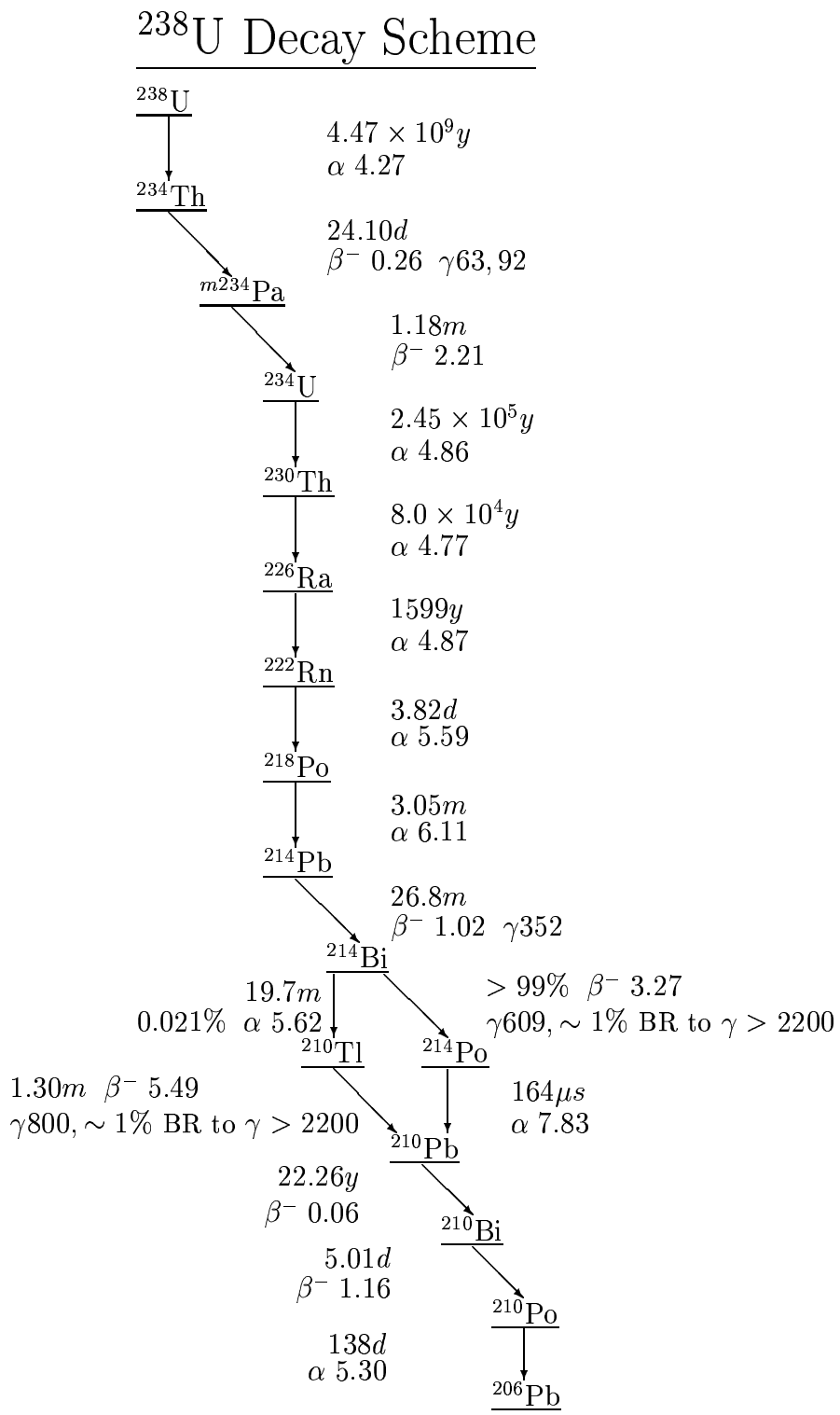


Figure A.2: The decay chain of ^{238}U. All half-lives are shown, with the Q-values of beta and alpha decays in MeV, and gamma rays in keV. Taken from [116].

Appendix B

Run Selection

It is necessary to collect solar neutrino data over long periods of time (many months) in order to acquire a sufficient sample of candidate neutrino events. During the salt phase of SNO, which lasted for 772 days, the detector was operated nearly continually. Approximately 48 days of data taking were lost due to technical problems such as power outages, broken fibre optical cables, and industrial action at the mine hosting the SNO experiment.

A large proportion ($\approx 20\%$) of time is devoted to calibrating the detector compared to a total of 525 days ($\approx 68\%$) spent collecting neutrino data [117]. Not all of the neutrino runs collected were suitable for solar neutrino analysis. To ensure that only good quality data was analysed, each run was subject to a series of criteria as detailed in this appendix. A series of tools were developed by the *Run Selection Committee* to apply these criteria to the candidate runs and generate a *run list* of quality neutrino runs for analysis. The author was a member of this committee throughout the course of salt data taking.

B.1 Selection Tools

Two SNOMAN processors were developed to extract the necessary parameters for each collected run:

- First Pass Run Selection (FSP)
  This processor makes tests of constants stored in the data stream for every run (including calibration and other non-neutrino type runs). Firstly a *bit-mask* defining the type of run is extracted. The detector operator* can alter this mask when starting a new run. For example, an *Unusual Circumstances* (UC) bit can be set to alert analysers to unusual run conditions such as when groups of visitors are touring

*The detector operator is responsible for making a number of checks during data taking and is present to take action if problems occur.

the deck area above the experiment. A test is also made on the types of trigger activated for each run and the status of all channels in the electronics (see 2.1.1 for a description). The duration of the run is also obtained by the FPS tool.

- Second Pass Run Selection (SPS)
  This processor makes more detailed checks of the data stored in a run and is only applied to neutrino-type runs. A large proportion of this code was written and developed by the author. In addition to a number of tests on the stored event times, the frequency of hits on individual channels, and on the separate triggers are measured. Errors output by electronics components are also analysed.

In addition to these automated tools, the run selection process involved analysing information stored in *shift reports* written by the detector operator for each shift, *builder log files* (see section 2.1.1 for a description of the *builder*) and *CMA log files*. The CMA is a system which monitors the environment in the SNO laboratory. The log files store information on temperature and humidity in the laboratory, and the status of equipment such as the compensation coils, which are described below.

B.2 Selection Criteria

The following criteria must be met in order for a run to be placed on the *Salt Run List*:

1. **Run Type**
   Only runs set as neutrino-type were selected. Runs with the UC (unusual circumstances) bit set as well were only selected if the reason for this was clearly documented and did not affect the data quality. Therefore, all runs with calibration sources inside the detector, or open gate-valves (through which the sources are deployed) were rejected.

2. **Run Length**
   Only runs greater than 30 minutes in length were accepted onto the run list as very short runs are usually stopped because of electronic problems. The run time was obtained from the output of FPS, from both the 10 MHz and 50 MHz clock. Agreement between these two times was also required. Runs were rejected if more than 20% of the run time was cut by data cleaning.

3. **10 MHz Clock Problems**
   The event times obtained from the 10 MHz clock were found to be faulty for a

number of runs. The times in these runs were corrected before the neutrino data was analysed.

4. **Deck Activity**
   Whilst human access to the deck area above the SNO detector does not normally affect data quality, disruptive work can cause electronic pick-up events which are hard to remove from the data set. Data collected when such work was being carried out were rejected.

5. **D_2O Circulation**
   The D_2O is frequently circulated in the detector to maintain the required levels of water purity. This activity is noted by the detector operator and stored in the bit-mask for each run. D_2O circulation does not appear to produce light, except when it is switched on and off. Therefore, only runs in which "circulation light" was observed by the operator were rejected.

6. **Bubblers**
   Occasionally air is bubbled through the D_2O in order to determine the water level, which can cause a lot of light in the detector. The *bubblers* must be off and flooded with water for a run to be accepted onto the run list.

7. **Assays**
   Only one type of assay, that samples D_2O from inside the AV, is known to produce light in the SNO detector. Details of such assays were flagged in the data, and the runs in which they were carried out were rejected.

8. **Compensation Coils**
   A number of *compensation coils* are installed above the SNO detector to correct for the affects of the Earth's magnetic field. Magnetic fields bend the tracks of electrons as they travel down the PMT dynode stacks. Cancelling the effect of this field increases the gain of the PMTs. For quality neutrino data taking all compensation coils must be operating at their nominal voltages.

9. **Temperature**
   The SNO laboratory is maintained at a constant temperature of 19°, but if part of the air-circulation system is turned off (for example during a short power-outage), the underground laboratory temperature can rise. High temperatures can have adverse affects on the electronics so runs were rejected if the recorded temperature (obtained from the CMA log files) was above 21°.

PMT type	Minimum number online
all PMTs	8600
OWLs and BUTTs	60
Necks	4

Table B.1: Minimum number of online PMTs.

10. **Event Rate**
    Event rates were checked by the SPS code and runs were rejected if the average rate was greater than 60 Hz, or if more than 20% of the run exceeded this threshold. Runs were also rejected if the rate of the pulsed global trigger deviated by more than 2% from the expected rate of 5 Hz.

11. **Invalid Electronics Calibration Constants**
    Turning off the high-voltage to the SNO detector can change the required calibration constants for the electronics. Only runs with valid Electronics Calibration (ECA) constants can be used for neutrino data analysis. Neutrino runs obtained immediately after a detector *shut-down* only passed the selection criteria if the ECA constants obtained from the next calibration were consistent with those before the shut-down.

12. **Trigger Thresholds**
    If trigger thresholds are vastly different to their normal values this can compromise data quality. The FPS code checks that the NHIT100 trigger (described in 2.1.1) threshold was below 20 hits. The nominal threshold for this trigger is 16 hits so if this check failed the run was rejected.

13. **Off-line Electronics**
    All crates must be operational for a good neutrino run, with no gaps in PMT coverage comparable in size, or larger than, the neck region of the detector. However, at present no automated check for this criterion has been written and selection relies on details obtained from *shift reports*. The FPS processor checks the number of PMTs online, and runs were rejected if the thresholds in table B.1 were not met.

    It is possible for crates or channels to trip off-line during a run, which can go unnoticed by the detector operator. Channel *occupancy*, the number of times a given channel fires, is measured over 20 minute integration periods by the SPS processor. The occupancy test will fail for a given integration period if the number of PMTs with zero occupancy exceeds the thresholds given in table B.2. If the amount of

PMT type	Maximum at zero occupancy
Normal PMTs	15
OWLs	4
BUTTs	4
Necks	0

Table B.2: Maximum number of zero occupancy PMTs.

integration periods for which the SPS occupancy check failed constituted more than 10% of a run it was rejected.

The PMTs installed in the neck region of the detector are necessary to reject a certain class of background events originating in this region. Therefore, a test is carried out to ensure that these PMTs are working. If more than one of these "Neck" PMTs failed the occupancy check for more than 10% of a run, it was rejected. A similar test was applied to outward-looking (OWL) PMTs which are necessary for rejecting events caused by muons.

14. **Blind Flashers**
One class of instrumental background, called *Flashers*, occur frequently (typically 50 per hour [118]) in the SNO detector. These events occur when a flash of light emitted by one PMT is detected by other PMTs and are easily rejected from the data set due to characteristic distributions in time and charge. However, if the signal from the PMT that emitted the flash of light is not read out, these events are not as easily distinguished. Candidates for such events, dubbed *blind flashers*, are PMTs which are supplied with a high voltage source, but have had their readout disabled. If there were more than eight possible blind flasher sources in a run, it was rejected.

15. **ESUM Trigger Occupancy**
This test, applied by the SPS processor, was designed to test the rate and occupancy of the ESUM trigger. If ESUM triggers are not received from all crates, the signal on the analogue measurement board (AMB) would be wrong. This is a potentially serious problem as the AMB signal is used to cut regular flasher background events. The ESUM trigger rate must be greater than 1 Hz, and a test is made to ensure that the fraction of events producing ESUM triggers is not significantly lower ($<$ 30% of the average over all crates) for any one crate.

16. **Sync Clear Errors**

    A test was made for errors returned by the CGT (Counter Global Trigger) electronics chip, which resets after 2^{16} events. If the CGT does not count 2^{16} events at the same time as the MTC/D it will synchronise to the MTC/D signal and set a *Sync clear* error flag. When the MTC/D uses up all of its 2^{24} GTIDs it wraps back round to GTID=1 and sends a *sync* signal to all cards in the electronics. If this sync signal comes when the CGT is not expecting it, a second error flag is set.

    If a channel is out of sync, PMT hits will be assigned to the wrong event, thus changing the *Nhits* of both the event it should have been assigned to and the event it was assigned to. The SPS code checks for CGT sync clear errors in the data stream, and counts the number of sync clear errors for each set of 2^{16} events. Since the energy scale uncertainty is $\approx$1%, fewer than 1% of boards were allowed to be out of sync for the length of the entire run. If more boards were out of sync for a small portion of the run, but on average accounted for less than 1% over the whole run, then the run was not rejected from the run-list.

17. **Sodium Activation**

    The data obtained following the deployment of calibration sources, such as the ^{16}N and ^{252}Cf sources, is subject to high levels of background from the decay of ^{24}Na. This isotope has a half life of 14.96 hours and is created by neutron capture on ^{23}Na. If the calculated number of neutrons produced by ^{24}Na decays in a run was greater than 5% of the SSM neutron flux, integrated over the run, then the run was rejected. If the run was not rejected, the number of background neutrons was calculated and included in the number of residual ^{24}Na neutrons.

18. **Salt Addition**

    There were a number of uncertainties about data quality immediately after the addition of salt to the detector. These included the levels of impurity in the water and the isotropy of the salt in the D_2O. Calibration studies were used to determine when the detector had stabilised and all data collected before run 20674 was removed from the runlist.

19. **Blasting**

    Runs are not normally rejected due to blasting in the mine. However, after large blasts or seismic activity in the mine, unusual electronic effects are occasionally observed. If undesirable effects were reported, the run was rejected from the run list.

Appendix C

Event Selection Criteria

In order to reduce the contributions of background events to the salt data set, a series of selection criteria ("cuts") were applied. This appendix details the criteria applied to both data and simulated events.

The selection criteria applied to the salt data are summarised in table C.1. The first entry in this table is the *DAMN mask* which characterises a series of cuts developed to remove instrumental backgrounds. It was applied as a bit test as described in Ref. [66].

The fiducial volume was restricted to radii $< 550\,\mathrm{cm}$ for two reasons. Firstly, to reduce the number of background events in the data set due to radioactivity in the PMT array, light water, and surrounding rock. Secondly, the detector response is less well understood at large radii because the acceptance of the PMT concentrators drops off at large angles from the normal. The kinetic energy threshold, $T_{eff} > 5.5\,\mathrm{MeV}$, also serves to remove low energy background events from the data sample.

The In-Time-Ratio (ITR) is a post-reconstruction test on the timing of an event, developed to reject poorly reconstructed Čerenkov events and instrumental backgrounds. It is the ratio of the number of hits, N_{it}, falling within an asymmetric window about the prompt event time, with respect to the total number of hits in the event [72].

The *muon follower cut* removes muon spallation products from the data stream. Neutrons produced after the passage of a muon through the SNO detector would result in a serious background to the NC signal. Events within 20 s of an identified muon event were cut from the data set and the livetime was corrected accordingly.

The selected range for the β_{14} isotropy parameters helped to limit the number of mis-reconstructed events in the data set, including backgrounds from outside the D_2O which reconstruct at low radii.

The criteria for event radius, and energy were also applied to simulated events. The DAMN mask and muon follower cut were not necessary for simulated events as instrumental backgrounds and muons were not included in the Monte Carlo of neutrino

Cut	**Parameters**
DAMN Mask	0x20f56fe1
Radius	$\leq 550\,\mathrm{cm}$
T_{eff}	$5.5 \leq T_{eff} < 20.0\,\mathrm{MeV}$
ITR	≥ 0.55
Muon Follower Cut	20 s following event > 60 hits
β_{14}	-0.12$\leq \beta_{14} < 0.95$

Table C.1: Event selection criteria applied to the salt data.

signals. The limits applied to β_{14} and ITR, are classed as *high level cuts* and were not applied to simulated events. This is because the number of good neutrino events excluded by these cuts, and the number of background events satisfying them, was calculated from comparisons of calibration data and simulated events. The correction calculated to account for these effects on the predicted number of events assumed that no high level cuts were applied to simulated events.

Appendix D

Isotropy Parameterisation

Due to the neutron capture mechanism on chlorine, NC events in the salt phase are more likely to produce multiple gammas, which give a wider distribution of Čerenkov light than the single electron produced in the CC and ES events. The isotropy of an event can be characterised in a number of ways as detailed below. A detailed discussion of studies related to the choice of isotropy parameter can be found in Ref. [57], but this appendix presents a small scale investigation into the merits of the chosen parameter, β_{14}.

D.1 Parameter Definitions

- $\boldsymbol{\theta}_{ij}$ - This parameter is defined as the average of the angles between each pair of PMT hits in an event with respect to the fitted vertex position.

- **Correlation Function Ring Inner Product - CFRIP** - This parameter [119] compares the *angular correlation function* of an event to that of a perfect ring of half angle 42° and width 8° using an *inner product* computation. The angular correlation function is essentially the distribution of cosines of angles between all possible pairs of hit PMTs in the event with respect to the fitted event vertex. If ϵ_i is the component of the angular correlation function in bin i for the tested event, and ρ_i is the component of the angular correlation function in the same bin for the ideal ring, the inner product is defined as:

$$\text{CFRIP} = \sum_{i=1}^{100} \epsilon_i \rho_i \tag{D.1}$$

  where data from the event was divided into 100 bins for ϵ_i in the range -1 to +1. This parameter provides good separation of CC and NC events as, on average, the more isotropic neutron events deviate more from the ideal case than electron-type events.

Isotropy PDFs	Mean Uncertainty		Mean Bias	
	CC	NC	CC	NC
θ_{ij}	31.12	31.48	-0.107 ± 0.875	0.140 ± 0.972
CFRIP	30.05	30.29	-0.033 ± 0.814	0.022 ± 0.085
β_{14}	29.02	29.23	-0.068 ± 0.800	-0.009 ± 0.878
β_1	31.26	31.64	-0.047 ± 1.237	-0.109 ± 1.307
β_4	42.30	42.08	0.137 ± 0.725	-0.156 ± 0.751
1D β_1 and 1Dβ_4	27.18	27.20	0.194 ± 0.818	0.095 ± 0.808
2D $\beta_1 \times \beta_4$	28.78	28.90	0.023 ± 0.813	-0.199 ± 0.987

Table D.1: The mean fit uncertainty and bias for CC and NC fluxes using different isotropy parameters.

- **β Parameters**: The lth beta parameter, β_l, is defined as the average value of the Legendre polynomial, P_l, of the cosine of the angle between each pair of PMT hits in the event.

$$\beta_l = \langle P_l(\cos\theta_{ik}) \rangle_{i \neq k} \tag{D.2}$$

  Again, the angle is taken with respect to the fitted vertex position. The combination $\beta_{14} = \beta_1 + 4\beta_4$ was selected by the SNO collaboration for use in signal extraction due to the good separability it provides and the ease of parameterisation of the Gaussian-like distribution.

D.2 Separation Power

To study the relative separation power of the different isotropy parameters, a test was conducted on simulated data sets. 100 data samples of 300 CC and 300 NC events falling in the energy range $5.5 < T_{eff} < 6.5$ were created from half of the simulated neutrino events. The narrow range in energy was used to avoid complications associated with the energy dependence of the isotropy parameters. In addition to the isotropy parameters, PDFs in R^3 and $\cos\theta_\odot$ were used in the likelihood fit although these provide little additional information for the separation of CC and NC signals. Separate fits were carried out on the same 100 data sets with different isotropy parameters including different combinations of β_1 and β_4 as detailed in table D.1.

The mean statistical uncertainty presented in table D.1 can be used as an indication of the quality of separation of the CC and NC signals. Uncertainties obtained using the parameter, β_{14}, were smaller than those obtained with θ_{ij} or CFRIP, indicating that β_{14} provides more information for the separation of the two signals. The mean uncertainty

on the fitted parameters was slightly smaller when β_{14} was used over β_1 alone and was significantly better than β_4 alone. Using both β parameters in separate one dimensional PDFs resulted in a smaller uncertainty than β_{14} though this did not take into account any correlation between these parameters which could cause a bias in the fit results.* Combining β_1 and β_4 in a 2 dimensional PDF produced slightly smaller uncertainties than using β_{14} but this would be more complicated to implement in a full fit where the energy dependence of isotropy becomes important and must be included as another dimension in the PDF.

*No bias was seen in this case though a bias was observed when the fit was performed on data over a larger energy range.

Appendix E

Sensitivity of the Likelihood Fit to PDF Bin Sizes.

An investigation was carried out to test the effect of varying the PDF bin sizes on the result of an energy-constrained likelihood fit. The PDF factorisation used for these tests is specified in equation 5.23. Therefore, the results of fits presented here should be compared to the results for this factorisation given in table 5.5, which used the default bin sizes given in table 5.1. Changes were made to the binning of each observable distribution in turn.

The quality of each fit was measured in a number of ways. Parameter bias and fractional deviation are defined in equations 5.18 and 5.19 respectively. The mean uncertainty, $\bar{\sigma}(X)$, on the fitted parameters is also indicative of the accuracy of the fit. If the magnitude of this uncertainty increases with coarser binning of an observable, this indicates possible removal of useful information from the fit.

E.1 Variations in T_{eff} Binning

The first alternative binning scenario for T_{eff}, used 29 0.5 MeV bins instead of a large final bin in the region of reduced statistics above 13.5 MeV. The second scenario used 15 bins of 1 MeV width in the range 5.5–20.5 MeV. The results of these two fit scenarios are presented in tables E.1 and E.2 respectively. Neither of these alternative binning scenarios had a significant effect on the biases observed. The uncertainties obtained from the fit with coarser T_{eff} binning were consistently larger, whilst there was little change in the mean fitted uncertainty for the first scenario. This indicates that a better fit was obtained with 0.5 MeV bins.

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2028.9 ± 6.8	0.04 ± 0.06	0.001 ± 0.002	68.466 ± 0.093
ES	309.1 ± 2.9	0.03 ± 0.08	0.004 ± 0.007	29.165 ± 0.067
NC	2170.8 ± 7.7	-0.23 ± 0.08	-0.008 ± 0.003	77.140 ± 0.095
EN	159.1 ± 4.4	0.24 ± 0.09	0.074 ± 0.026	44.292 ± 0.081

Table E.1: Results for fits to 100 artificial data sets with 29 0.5 MeV T_{eff} bins.

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2019.8 ± 6.9	-0.09 ± 0.06	-0.003 ± 0.002	68.757 ± 0.095
ES	313.9 ± 2.9	0.19 ± 0.07	0.019 ± 0.007	29.320 ± 0.066
NC	2179.0 ± 7.8	-0.12 ± 0.08	-0.004 ± 0.003	77.803 ± 0.093
EN	156.4 ± 4.5	0.18 ± 0.09	0.055 ± 0.027	44.811 ± 0.079

Table E.2: Results for fits to 100 artificial data sets with 15 1 MeV T_{eff} bins.

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2019.0 ± 6.9	-0.10 ± 0.06	-0.003 ± 0.002	68.593 ± 0.094
ES	313.9 ± 2.9	0.19 ± 0.07	0.019 ± 0.007	29.315 ± 0.066
NC	2179.3 ± 7.8	-0.12 ± 0.08	-0.004 ± 0.003	77.668 ± 0.096
EN	156.9 ± 4.5	0.19 ± 0.09	0.058 ± 0.027	44.787 ± 0.077

Table E.3: Results for fits to 100 artificial data sets with 25 equal sized bins for the β_{14} distribution.

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2050.6 ± 6.8	0.36 ± 0.06	0.012 ± 0.002	68.315 ± 0.094
ES	304.7 ± 2.9	-0.13 ± 0.07	-0.010 ± 0.007	28.898 ± 0.067
NC	2146.0 ± 7.6	-0.56 ± 0.07	-0.019 ± 0.003	75.708 ± 0.092
EN	166.1 ± 4.3	0.42 ± 0.09	0.123 ± 0.025	42.684 ± 0.100

Table E.4: Results for fits to 100 artificial data sets with 20 equal sized bins for the R^3 distribution.

E.2 Variations in β_{14} Binning

To test the sensitivity to binning of the isotropy distribution, the number of β_{14} bins was reduced* from 50 to 25. The results of this fit are given in table E.3. Whilst this reduction in β_{14} bins did not introduce any significant bias to the fit, the mean fitted uncertainties were slightly larger, suggesting that useful information was removed from these PDFs, and the use of 50 β_{14} bins gives a better fit.

E.3 Variations in R^3 Binning

A small number of R^3 bins was selected in order to minimise the number of bins in the 3-dimensional PDF. To test that this coarse binning did not withold important information from the fit, signal extraction was performed on the 100 simulated data sets with PDFs created in 20 equal sized R^3 bins. The results of this fit are given in table E.4 and show significantly larger biases than the case with 10 R^3 bins. This effect is thought to be due to statistical fluctuations leading to systematic biases in the 3-dimensional PDF shape. In this scenario there were $17 \times 50 \times 20 = 17000$ bins in the three-dimensional PDF, and $100 \times 20 = 20000$ bins in the conditional two-dimensional PDF, but insufficient simulated events available to accurately represent the predicted distributions at this level.

E.4 Variations in $\cos\theta_\odot$ Binning

For good sensitivity to the sharp ES peak in the $\cos\theta_\odot$ distribution, 100 bins were used as default for this distribution. To test the sensitivity of the fitting procedure to the information in this PDF, the fit was repeated with only 25 $\cos\theta_\odot$ bins. The results of this test are given in table E.5.

*An increase in the number of β_{14} bins could not be easily tested due to limitations on the size of arrays containing multi-dimensional PDFs.

X	$\bar{F}(X)$	Mean $Bias(X)$	$\bar{\Delta}(X)$	$\bar{\sigma}(X)$
CC	2028.8 ± 6.9	0.04 ± 0.06	0.001 ± 0.002	68.540 ± 0.093
ES	310.8 ± 2.9	0.08 ± 0.07	0.009 ± 0.007	29.478 ± 0.066
NC	2171.0 ± 7.7	-0.23 ± 0.08	-0.008 ± 0.003	77.212 ± 0.095
EN	158.3 ± 4.4	0.22 ± 0.09	0.068 ± 0.026	44.370 ± 0.081

Table E.5: Results for fits to 100 artificial data sets with 25 equal sized bins for $\cos\theta_{\odot}$.

Unsurprisingly, the change in $\cos\theta_{\odot}$ binning did not introduce any significant bias to the analysis, and only affected the size of the fitted uncertainties on the CC and ES parameters (the neutron distributions are flat in $\cos\theta_{\odot}$).

E.5 Discussion

The binning tests for R^3 indicate problems associated with limited statistics in multi-dimensional PDFs. However, there is no evidence for biases introduced by reducing the number of PDF bins. Whilst the effects of coarser PDF binning are noticeable in the magnitude of the fitted uncertainties for all parameters, the size of this effect is very small, in all cases less than 1% of the fitted uncertainties. Such effects are insignificant compared to the contributions of systematic uncertainties in the observable distributions that are discussed in detail in section 6.2.

Appendix F

Aborted Simulation Events

The SNOMAN simulation package has been well written and tested against all available calibration data but the complexity of the code is such that errors due to rare occurrences have not been entirely eliminated. An example is the tracking of photons through unusual detector regions, which cause SNOMAN to abort a single event in every few hundred generated. The number of these "geometry errors" is expected to be larger for higher energy events as more photons are generated. Errors in the Monte Carlo Data Acquisition (MCDAQ) routine also occur very infrequently. The occurrence of such errors can artificially reduce the number of events in a Monte Carlo data set. As the Monte Carlo is used to estimate the number of events inside the analysis cuts for each neutrino type, corrections for these errors are required. An investigation carried out to determine a suitable energy-dependent correction is presented here.

F.1 Method

Separate Monte Carlo simulations of mono-energetic electrons were performed for 1 MeV steps over the kinetic energy range 4–20 MeV, and in 5 MeV steps from 20–60 MeV.* For each of these simulations, the same optical settings were used as for salt data processing and run conditions were kept the same throughout for simplicity.† Samples of 200000 events were simulated at each electron energy.

*The higher energy measurements were carried out for use in analysis of the *hep* signal in SNO.

†Run conditions, including the number and position of on-line tubes, for run 26997 were used throughout. The alternative would be to weight the simulation over the entire salt data set, which would require significantly more book-keeping. This is a second order effect to a small correction.

F.2 Results

Table F.1 summarises the number of geometry and MCDAQ errors that occurred at each generated energy. The correction factors required to convert from the number of Monte Carlo events generated to the true number of events predicted (shown in figure F.1) can be parameterised by

$$MC_{\text{corr}} = 0.99995(\pm 0.67 \times 10^{-4}) - 0.0006845(\pm 0.39 \times 10^{-5}) \times E_{\text{gen}}. \qquad \text{(F.1)}$$

where E_{gen} is the total generated electron energy in MeV, and uncertainties on the fit parameters are given in parentheses. Although the difference between the generated electron energy and the observed energy T_{eff} may differ slightly from the electron mass energy, this is a small second order effect, and $\frac{1}{MC_{\text{corr}}}$ was used to weight generated events based on their value of $(T_{eff} + 0.511)$ MeV, rather than E_{gen}.

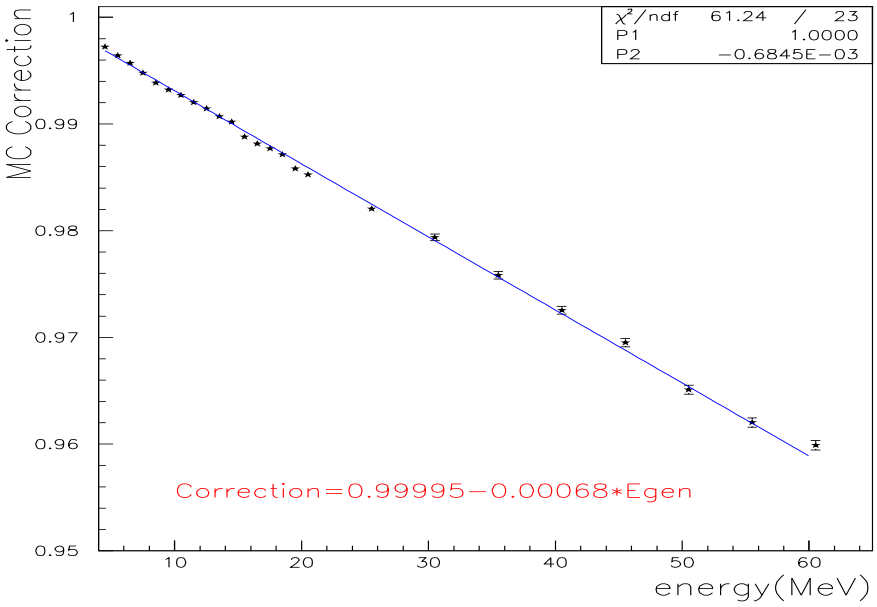


Figure F.1: Correction factor required to convert from the number of Monte Carlo events generated to the true number of events predicted, compensating for aborted simulation events.

E_{gen} (MeV)	MCDAQ	GEOM	MC_{corr}	$\delta(MC_{\text{corr}})$
4.511	2	553	0.99723	0.00012
5.511	8	712	0.99640	0.00013
6.511	4	855	0.99570	0.00015
7.511	3	1039	0.99479	0.00016
8.511	2	1229	0.99384	0.00018
9.511	2	1358	0.99320	0.00018
10.511	3	1454	0.99272	0.00019
11.511	3	1593	0.99202	0.00020
12.511	2	1709	0.99145	0.00021
13.511	2	1859	0.99069	0.00022
14.511	2	1959	0.99020	0.00022
15.511	0	2240	0.98880	0.00024
16.511	0	2369	0.98816	0.00024
17.511	1	2461	0.98769	0.00025
18.511	2	2569	0.98715	0.00025
19.511	1	2834	0.98582	0.00027
20.511	1	2946	0.98526	0.00027
25.511	2	3588	0.98205	0.00030
30.511	3	4121	0.97938	0.00032
35.511	2	4833	0.97583	0.00035
40.511	2	5490	0.97254	0.00037
45.511	1	6096	0.96952	0.00039
50.511	1	6979	0.96510	0.00042
55.511	1	7593	0.96203	0.00044
60.511	0	8022	0.95989	0.00045

Table F.1: Monte Carlo corrections due to MCDAQ and geometry errors for different electron energies. The columns give the total electron energy generated, the number of MCDAQ errors, and the number of geometry errors occurring. The number of simulated events should be divided by MC_{corr} to obtain the true number of events.

Appendix G

Energy Non-linearity

The relationship between deposited and observed energy (*energy-mapping*) in the SNO detector was predicted to be nearly linear by both Monte Carlo simulations and the analytic calculation presented in chapter 3. Whilst the agreement between data and Monte Carlo has been well studied at ^{16}N energies, non-linear effects omitted from (or poorly modelled in) the Monte Carlo simulations could cause systematic differences between the data and simulations at higher energies. Studies carried out to investigate the magnitude of this systematic effect are presented here.

G.1 Previous Studies

In the D_2O phase the proton-tritium (Pt) calibration source, which produces 19.8 MeV γs was used to place limits on deviations from linear energy-mapping. The good agreement of the observed and predicted mean T_{eff} value for this source, along with independent calculation of the magnitude of cross-talk and multi-photon effects at high energy, was used to limit the non-linearity uncertainty to 0.25% for the Phase I data. The agreement of the energy calculated in chapter 3 with Pt calibration data at the 1% level also provided a good argument against non-linear effects in the Phase I data. The same study could not be performed to limit the non-linearity uncertainty of the salt data as no Pt calibration data was obtained in Phase II.

Addition of NaCl to the D_2O for Phase II changed the optical properties of the SNO detector, but these changes have been well calibrated. The two known mechanisms that could affect the energy-mapping in a non-linear manner are electronic cross-talk/pickup and multiple photo-electron effects. As there is no obvious mechanism whereby the addition of salt could change these electronic effects, it can be argued that the same uncertainty on energy non-linearity should apply to the Phase II data. This argument is supported by studies of ^{8}Li data collected in both phases that resulted in consistent models for the

cross-talk phenomenon [120, 121] and a study that showed incorrect modelling of multiple photons to have only a small effect on the predicted energy of events ($\leq 0.6\%$ at 18 MeV [122]).

However, it is possible that other unknown phenomena could affect the energy-mapping at high energies. The approach adopted for this thesis, was to use calibration data to place limits on the difference between observed and predicted T_{eff} values across the energy range of interest in order to determine an estimate for the non-linearity systematic uncertainty.

G.2 Method

The ^{8}Li source [123] was only deployed once in the SNO detector during the salt phase providing the data used for this analysis, which is summarised in table G.1. In the following discussion, the set of runs at $R_s = 21.59$ cm are called the "central" runs.

The ^{8}Li source was used to study differences between data and simulated events in the T_{eff} range 4–15 MeV, whilst the ^{16}N source was used to provide an accurate constraint on the differences at the ^{16}N energy. ^{16}N calibration runs were chosen at similar locations to the ^{8}Li data for the purposes of this study [124]. Data obtained in the D_2O phase with the Pt source was also used to constrain differences between data and simulation at $T_{eff} \approx 18.5$ MeV.

The ^{16}N "reference energy" was taken as the mean T_{eff} value obtained from a fit to data for each run used in the T_{eff} range 3.5–6.5 MeV. The difference between data and Monte Carlo simulated events was quantified by the weighted mean $\bar{\Delta}(^{16}\mathrm{N})$ of

$$\Delta(^{16}\mathrm{N})_{\mathrm{run}} = \frac{F_{\mathrm{data}}(T_{eff}) - F_{\mathrm{MC}}(T_{eff})}{F_{\mathrm{data}}(T_{eff})} \tag{G.1}$$

in which $F_{\mathrm{data}}(T_{eff})$ and $F_{\mathrm{MC}}(T_{eff})$ are the fitted peak T_{eff} values for data and simulated events, respectively.

Table G.2 summarises the selection criteria applied to ^{8}Li data and simulated events for this analysis. The ^{8}Li data was processed as described in Ref. [125]. For ^{8}Li there is not a fixed event energy, but a spectrum similar to that for CC events. The relative difference between data and simulation as a function of energy was obtained from the normalised spectra in 1 MeV wide bins. For the ith bin, the difference was quantified by

$$\Delta(^8\mathrm{Li})_i = \frac{A_{\mathrm{data}}(i) - A_{\mathrm{MC}}(i)}{A_{\mathrm{data}}(i)} \tag{G.2}$$

in which $A_{\mathrm{data}}(i)$ and $A_{\mathrm{MC}}(i)$ are the amplitude of the data and simulated energy spectra in bin i, respectively.

	Pivot position (cm)			Source Position (cm)
Run number	x	y	z	R_s
23108	0	-21.59	85.49	21.59
23058	0	-21.59	85.58	21.59
23056	0	-21.59	85.58	21.59
23054	0	-21.59	85.56	21.59
23052	0	-21.59	85.52	21.59
23050	0	-21.59	85.57	21.59
23105	0	-21.59	335.57	251.1
23104	0	-21.59	335.57	251.1
23103	0	-21.59	335.57	251.1
23077	0	-21.59	-168.76	255.1
23075	0	-21.59	-168.76	255.1
23073	0	-21.59	-168.77	255.1
23091	384.93	-1.32	159.62	392.0
23088	384.92	-1.32	159.6	392.0
23086	384.91	-1.31	159.56	392.0
23094	0	-21.59	535.69	450.8

Table G.1: Run numbers and source positions for the ^{8}Li data. The x, y and z coordinates give the position of the manipulator 'pivot' which is 85 cm above the center of the ^{8}Li source. R_s is the radial position of the centre of the ^{8}Li source.

Parameter	Cut
β_{14}	-0.12 $< \beta_{14} <$ 0.95
ITR	ITR > 0.55
R	$0 < R <$ 850 cm

Table G.2: Event selection criteria applied to ^{8}Li data and simulated events. ITR is the in-time-ratio defined in appendix C and R is the reconstructed event radial position.

R_s	Mean T_{eff} Peak	Δ(Pt)
center	18.8 $\pm$ 0.1	0.5 $\pm$ 0.3%
500 cm	18.5 $\pm$ 0.1	-1.2 $\pm$ 0.6%

Table G.3: Values for the mean peak energy of the Pt source and the percentage difference between data and simulations for different radial positions of the source. Values taken from Ref [126].

The values used for the Pt source are given in table G.3 and were obtained from Ref [126]. Δ(Pt) is defined as for $\Delta(^{16}\mathrm{N})$. Since this data was obtained under different detector conditions, an additional uncertainty accounting for changes in the known non-linear effects should be included in Δ(Pt). However, separate studies of multi-photon effects and cross-talk have indicated that this additional uncertainty is negligible.

$\bar{\Delta}(^{16}\mathrm{N})$, Δ(Pt) and the eleven $\Delta(^{8}\mathrm{Li})_i$ values defined above were plotted together for each ^{8}Li source position. Care was taken to match runs with similar source positions for the ^{16}N and Pt source (all off-centre source positions assume the larger uncertainty obtained for the Pt source at 500 cm). Linear and quadratic fits to the data at each source position are shown in figure G.1. The fit parameters and the χ^2 for each fit is given in table G.4. For the center, R_s=255 cm and R_s=450 cm runs, the χ^2 per degree of freedom is better for the linear fit than the quadratic fit. The other two source positions were located on the positive z axis of the detector where the geometry effects are less well understood due to the neck region of the detector.

G.3 Results

The fitted slopes for each source position are plotted together in figure G.2 (solid lines) along with 1σ uncertainties (dashed lines). The largest positive slope is taken as a conservative limit on deviations from linear scaling in energy, as shown by the thick dashed lines on this figure. This results in an uncertainty of:

$$\sigma_{\mathrm{nl}} = 0.09\ \%\ /\ \mathrm{MeV}.$$

G.4 Application to Signal Extraction

To evaluate the effect of this uncertainty, the maximum likelihood fit was repeated with the value of T_{eff} for each event contributing to the PDFs scaled by the factor:

$$\omega_{\mathrm{nl}} = 1.0 \pm 0.0009(T_{eff} - 5.05). \qquad \text{(G.3)}$$

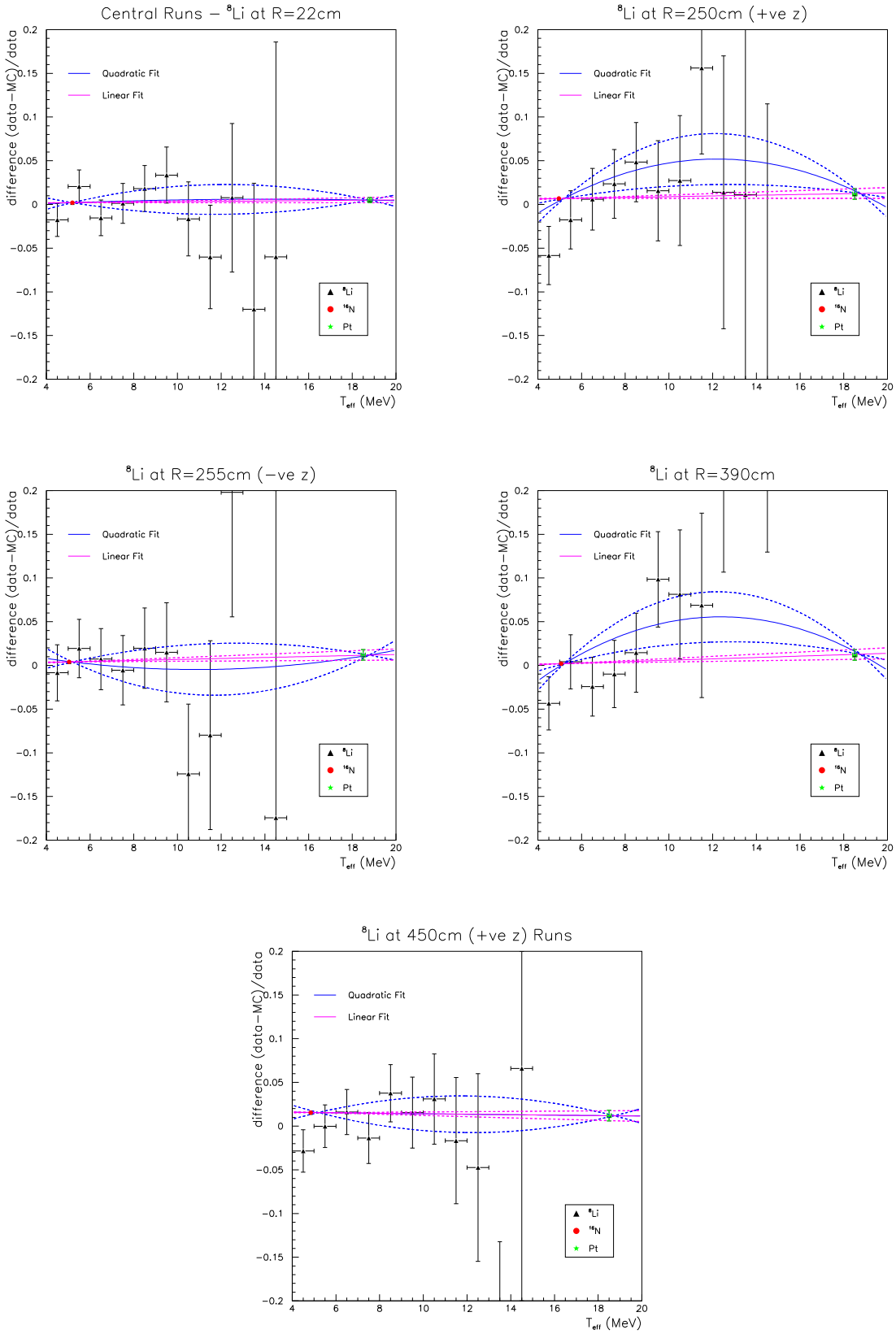


Figure G.1: Linear and quadratic fits (solid lines) to the differences between data and simulation for ^{8}Li, ^{16}N and Pt at five different source positions. The $\pm 1\sigma$ uncertainty bands on each fit are also shown (dashed lines).

	$\Delta = a + bx$ (10 DoF)			$\Delta = a + bx + cx^2$ (9 DoF)			
R_s	χ^2	a	b	χ^2	a	b	c
22 cm	6.22324	0.00098 ±0.00116	0.00022 ±0.00022	6.20663	-0.00381 ±0.03609	0.00138 ±0.00887	-0.00005 ± 0.00037
251 cm	8.26436	0.00458 ±0.00252	0.00043 ±0.00044	6.15911	-0.08494 ±0.06185	0.02242 ±0.01519	-0.00092 ±0.00063
255 cm	9.92573	0.00094 ±0.00238	0.00058 ±0.00044	9.75197	0.02694 ±0.06274	-0.00581 ±0.01541	0.00027 ±0.00064
392 cm	15.8536	-0.00226 ±0.00242	0.00080 ±0.00044	13.06068	-0.10416 ±0.06094	0.02585 ±0.01497	-0.00105 ±0.00062
450 cm	8.71465	0.01689 ±0.00242	-0.00027 ±0.00044	8.71460	0.01721 ±0.04455	-0.00035 ±0.01094	0.000003 ±0.00046

Table G.4: χ^2 values and fitted parameters for linear and quadratic fits to ^{8}Li, ^{16}N and Pt data-simulation fractional differences. Uncertainties are given below each fitted parameter. The number of degrees of freedom (DoF) for each fit is also given.

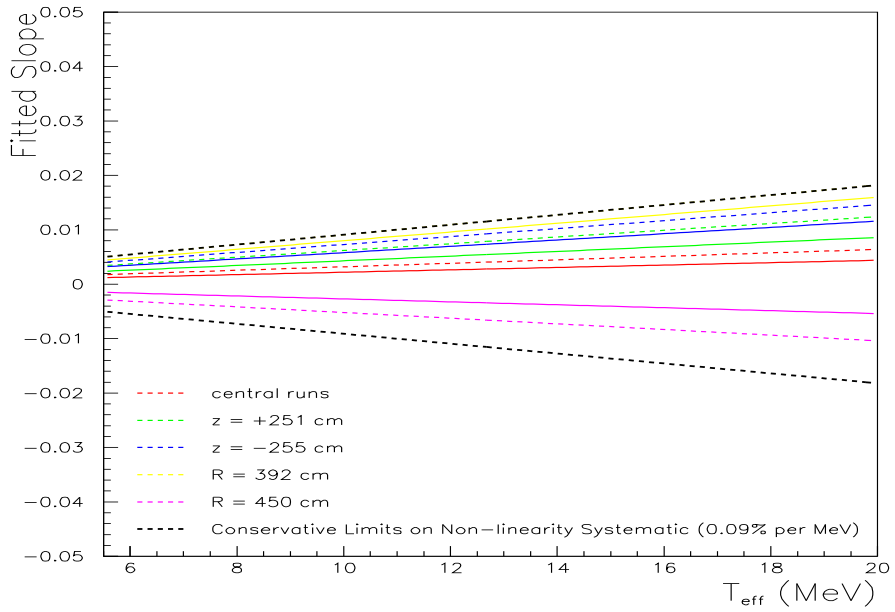


Figure G.2: Uncertainties defined by the gradients of the linear fits to ^{8}Li, ^{16}N and Pt data-simulation differences, plotted in G.1. The thicker black dashed line indicates a conservative limit to the deviation from linear scaling.

Appendix H

Energy Dependent Fiducial Volume Uncertainty

Biases in vertex reconstruction could pull events in to or out of the fiducial volume. Energy-dependent biases of this nature would affect the number of events inside the fiducial volume at different energies, thus distorting the spectrum. Ideally, calibration data obtained at $R \approx 550\,\mathrm{cm}$ (the edge of the fiducial volume selected for data analysis) spanning a range of energies, would be used to quantify this effect. In the absence of such ideal data, a number of approaches were investigated as detailed in Ref. [127], but a study comparing ^{8}Li calibration data with simulated events was found to be the most suitable and is presented here.

H.1 Method

Data obtained from the ^{8}Li calibration source (as detailed in table G.1) was used for this study as it samples a similar energy distribution to that of CC events. The event selection criteria detailed in table G.2 were applied to the data.

The effects of both vertex resolution and vertex shift with energy were studied simultaneously through the fraction of events that reconstruct inside the true vertex position, $N(R_{\mathrm{fit}} < R_{\mathrm{source}})$:

$$F_{\mathrm{in}} = \frac{N(R_{\mathrm{fit}} < R_{\mathrm{source}})}{N_{\mathrm{total}}} \tag{H.1}$$

where N_{total} is the total number of selected events. F_{in} should be approximately 0.5 but will be modified by solid angle effects that depend on source position. These effects should all be modelled in the Monte Carlo simulations, so a comparison of F_{in} between data and simulated ^{8}Li events, as given in equation H.2, should indicate any systematic effects with energy.

$$\Delta(F_{\rm in}) = F_{\rm in}({\rm data}) - F_{\rm in}({\rm MC}) \tag{H.2}$$

Consider multiple point sources on the boundary of the fiducial volume. If the value of $F_{\rm in}$ is the same for each of these sources, it can be used to estimate the number of events mis-reconstructed inside the fiducial volume. Furthermore, $\Delta(F_{\rm in})$ can be used to characterise the uncertainty in this number. The effects of additional point sources inside and outside the fiducial volume boundary will cancel provided they extend to sufficiently high radii and have the same values of $\Delta(F_{\rm in})$. As the vertex resolution in SNO is $\approx 16\,$cm and the D_2O volume extends to 600 cm, 50 cm beyond the fiducial volume boundary, this assumption should be valid. No ^{8}Li calibration data was available at a radius of 550 cm, the boundary of the fiducial volume, but as no strong position dependence in $\Delta(F_{\rm in})$ was observed, the values obtained from available data were assumed to apply at the boundary of the fiducial volume.

The source position was taken as the peak of the reconstructed radial distribution for both data and simulated events. Small uncertainties in this value would produce an overall offset in $\Delta(F_{\rm in})$, but would not contribute to any differential effect in energy. The analysis was carried out separately for each of the five ^{8}Li source positions. Nine T_{eff} bins in the range 3.5–12.5 MeV were used although the higher bins were severely statistically limited.

H.2 Results and Discussion

The value of $F_{\rm in}$ for both data (black triangles) and simulated events (red squares) is plotted against energy for each source position in figure H.1. This figure also shows a linear fit to $\Delta(F_{\rm in})$ for each position. The offset of this fit relates to an overall systematic in vertex reconstruction, but such effects were characterised much more accurately through studies of ^{16}N data (see section 6.2.4). The slope, however, corresponds to an energy-dependent systematic. In figure H.2 the fitted slopes from figure H.1 are shown as a function of source position.

The assumption that the vertex resolution is much less than the radius breaks down at low radii and solid angle effects skew the radial distribution. Therefore low radius runs may be more sensitive to source effects. For all but the lowest radius run, there is no evidence of a trend in fitted slope with source position. Studies of ^{16}N calibration data do not indicate any radial dependence in the reconstruction uncertainty either [95]. Therefore, to improve statistics the $F_{\rm in}$ values for all runs at $R_{\rm s} > 100\,$cm were combined, and fitted following the above prescription. The resultant fit is shown in in figure H.3.

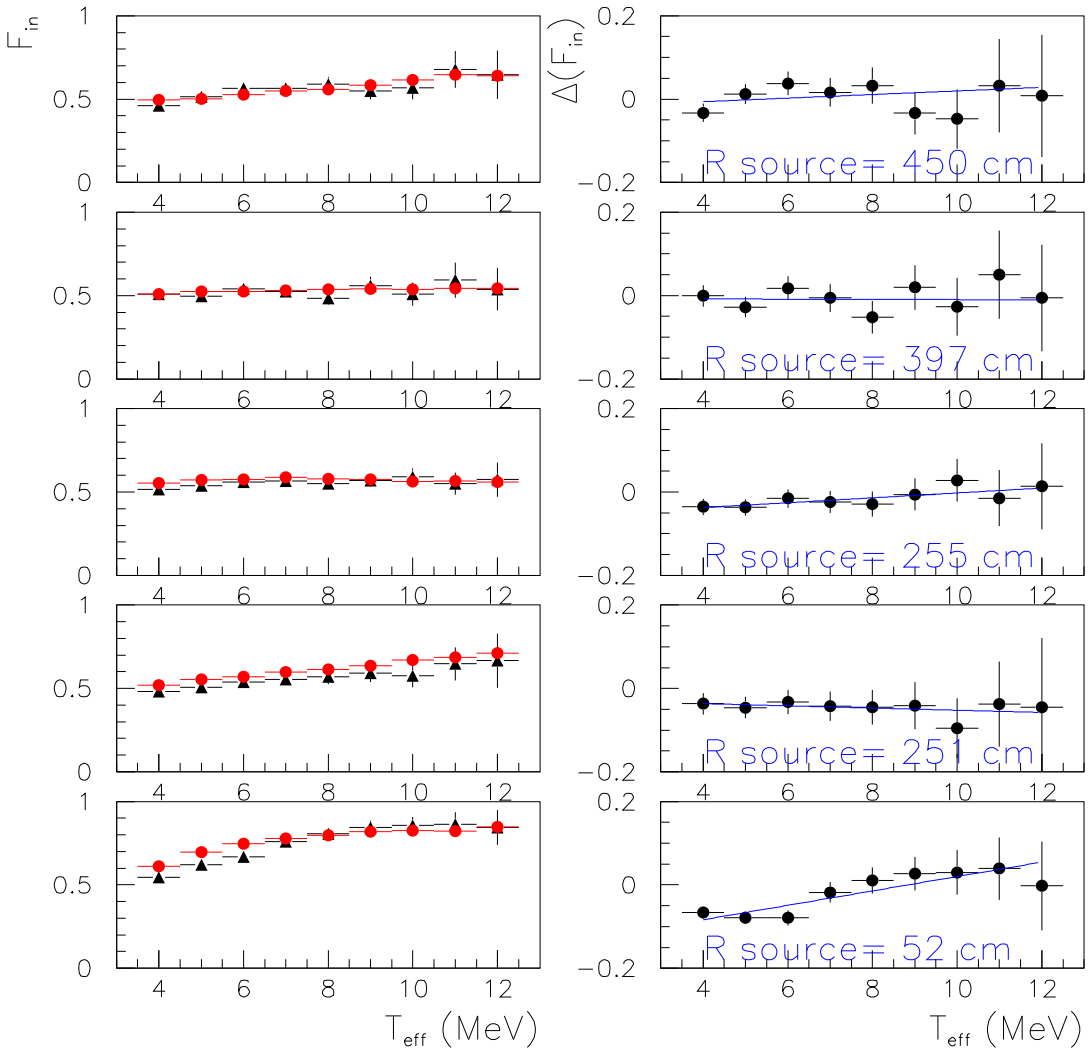


Figure H.1: The ratio of events fit with $R < R_s$ to total events (F_{in}) for data (black triangles) and simulated events (red squares) for all five ^{8}Li source positions. A linear fit to the difference between data and simulation ($\Delta(F_{in})$) is shown on the right hand side.

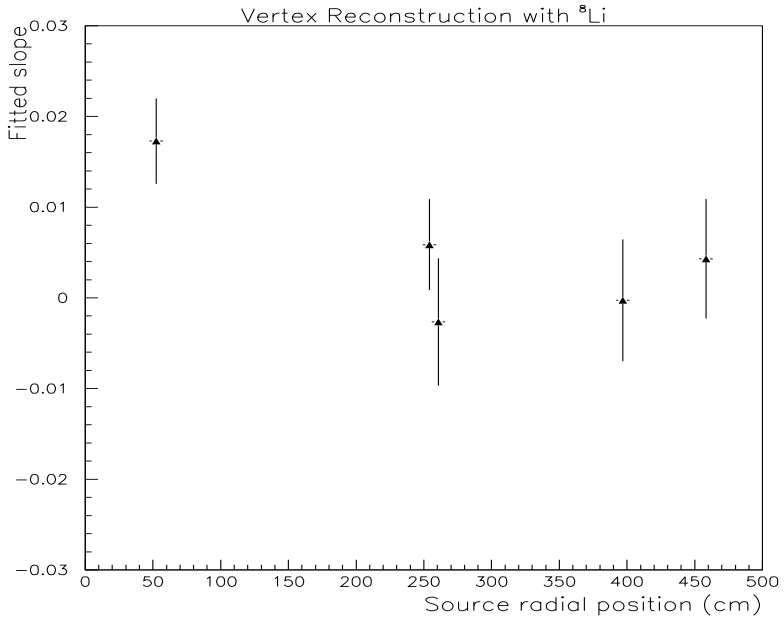


Figure H.2: Fitted relation between $\Delta(F_{\rm in})$ and T_{eff} plotted against source position obtained from ^{8}Li

The gradient of this fit was

$$\delta_{\rm F} = (0.32 \pm 0.31)\ \%\ /\ {\rm MeV} \tag{H.3}$$

which can be used to quantify the energy-dependent systematic in the number of events reconstructing inside the fiducial volume. As the main studies of vertex reconstruction systematic effects were obtained from the ^{16}N source, the energy-dependent systematic was constrained to zero at T_{eff}= 5.05 MeV, the mean observable energy of ^{16}N events. Thus, for example, the systematic difference at 13 MeV is $(13 - 5.05) \times \delta_{\rm F} = 2.54\%$

One possible caveat to this analysis is the fact that higher energy electrons travel further from the ^{8}Li source and may therefore be less sensitive to shadowing effects. To test the impact of this possible source effect, sets of ^{8}Li events were simulated with different models for the reflectivity of the exterior stainless steel shell of the source. The fitted values of $\delta_{\rm F}$ (given in table H.1) did not change significantly with source reflectivity.*

*Note that this study was carried out with lower statistics Monte Carlo than the simulations used for the rest of this study.

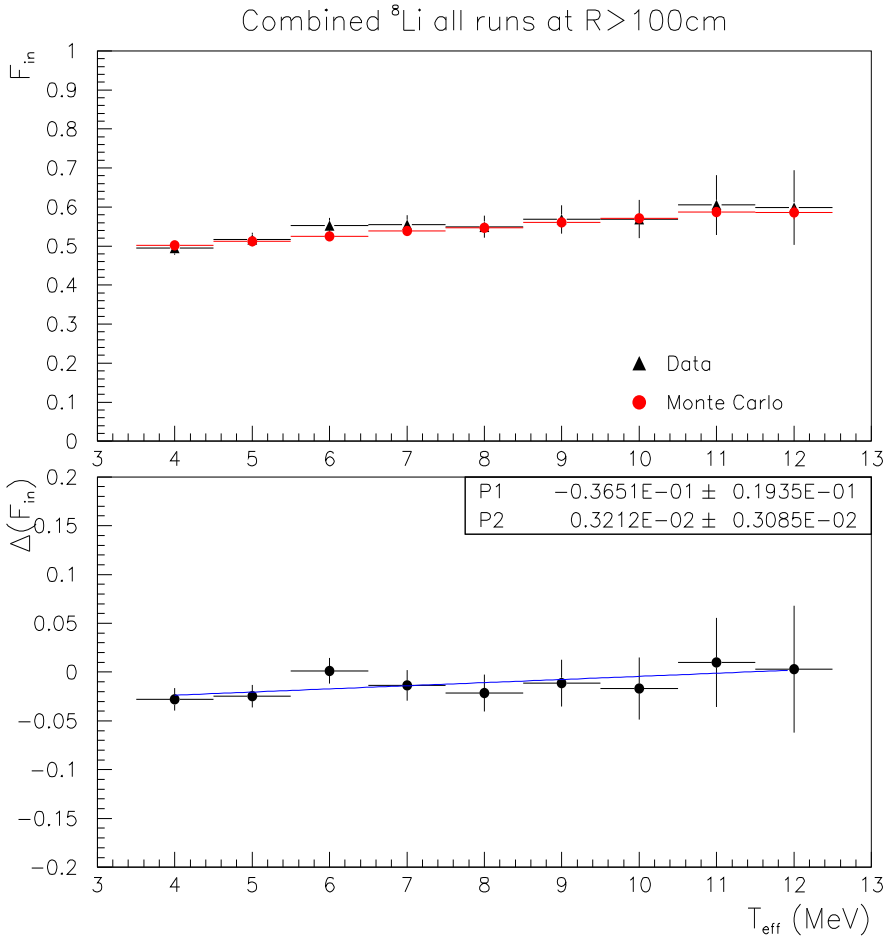


Figure H.3: The ratio of events inside the source position to total events for data (black) and simulated events (red) combined for ^{8}Li source positions with R>100 cm. The lower plot shows a linear fit to the difference between data and simulations against energy.

Simulation	Reflection probability	δ_F
Default	0.3	$(0.46 \pm 0.37)\%$
Black	0.0	$(0.32 \pm 0.37)\%$
Reflective	0.5	$(0.18 \pm 0.37)\%$

Table H.1: Gradient of fits to $\Delta(F_{in})$ versus T_{eff} when different models of source reflectivity were used in the Monte Carlo.

H.3 Application to Signal Extraction

The results of this study indicate that the ratio of data events to simulated events reconstructing inside the fiducial volume increases with energy, at a rate of $\delta_{\rm F}$ per MeV (as given in equation H.3). To correct for this effect, each simulated event contributing to a PDF was weighted by the factor

$$V_{\rm corr}(T_{eff}) = 1.0 + \delta_{\rm F}(T_{eff} - 5.05\,{\rm MeV}). \tag{H.4}$$

The effect of the uncertainty on this systematic was evaluated by using the $\pm 1\sigma$ limits on $\delta_{\rm F}$ in the PDF weighting factors for separate likelihood fits.

Appendix I

Predicted Fluxes

Monte Carlo simulation was used in order to predict the number of events due to each neutrino interaction in the salt data set. Interactions due to neutrinos produced by both ^{8}B and *hep* reactions were included in this prediction. A number of corrections were required to account for effects omitted from these simulations, as detailed in this appendix.

I.1 PDF corrections

The corrections applied are summarised in table I.1 and explained below. Whilst most corrections change only the predicted fluxes, some exhibit an energy dependence that can change the shape of PDFs used in the signal extraction process. The treatment of these effects is explained in chapter 6.

$\boldsymbol{R^{\mathrm{B}}_{\mathrm{model}}}$ **: Model flux correction for ^{8}B neutrinos.** The Monte Carlo simulation assumed a flux of $5.15\times10^6\,\mathrm{cm^{-2}\,s^{-1}}$ [128] for neutrinos produced by the ^{8}B reaction in the Sun. This correction factor was applied to convert to the BP2004 model prediction [21] of $5.82\times10^6\,\mathrm{cm^{-2}\,s^{-1}}$.

$\boldsymbol{R^{\mathrm{hep}}_{\mathrm{model}}}$ **: Model flux correction for *hep* neutrinos.** The Monte Carlo simulation assumed a flux of $9.24\times10^3\,\mathrm{cm^{-2}\,s^{-1}}$ [128] for neutrinos produced by the *hep* interaction in the Sun. This correction factor was used to convert to the BP2004 model prediction [21] of $7.88\times10^3\,\mathrm{cm^{-2}\,s^{-1}}$. The uncertainty on this newer model is $\pm16\%$, but no uncertainty was given on the earlier model prediction.

$\boldsymbol{\omega^{\mathrm{RC}}_{\mathrm{X}}}$ **: Radiative corrections.** Corrections due to the radiation of Bremsstrahlung photons could be included in the simulation but it would be time consuming to calculate these corrections on an event-by-event basis. Instead the weights given in equations I.1 and I.2 were applied to each event as it was assigned to a PDF. The CC

Correction	Energy Dependent	CC	ES	NC
$R^{\rm B}_{\rm model}$	No	1.130		
$R^{\rm hep}_{\rm model}$	No	0.853		
$\omega^{\rm RC}_{\rm X}$	Yes	equation I.1	equation I.2	$\omega^{\rm RC}_{\rm NC} = 1.0154$
$\alpha^{\rm B}_{\rm MC}$	No	200	200	-
$\alpha^{\rm hep}_{\rm MC}$	No	2000	2000	-
$N_{\rm D}$	No	1.01228	-	1.01228
$N_{\rm e}$	No	-	1.0151	-
$S_{\rm ac}$	Yes	See table I.2		
$R_{\rm O,Na,Cl}$	No	1.0081	-	-
$g_{\rm A}$	No	1.0111	-	1.0111
$L_{1,\rm A}$	No	0.984	-	0.979
$MC_{\rm err}$	Yes	See appendix F		
$\frac{t'}{t}$	No	$\frac{391.43}{398.59} = 0.9820$		

Table I.1: Corrections to number of events predicted by the signal Monte Carlo.

radiative correction is given in terms of the generated event energy, $E_{\rm gen}$ [107] whilst the ES radiative correction was determined empirically [129] in terms of the observed energy, T_{eff}.

$$\omega^{\rm RC}_{\rm CC} = 1.0318 - 7.45 \times 10^{-4} E_{\rm gen} + 4.72 \times 10^{-6} E^2_{\rm gen} \qquad \text{(I.1)}$$

$$\omega^{\rm RC}_{\rm ES} = 0.9764 - 0.781 \times 10^{-4} T_{eff} - 1.31 \times 10^{-4} T^2_{eff} + 3.64 \times 10^{-6} T^3_{eff} \qquad \text{(I.2)}$$

For NC events, the radiative correction factor was not energy-dependent and each event was weighted by $\omega^{\rm RC}_{\rm NC}$.

$\alpha_{\rm MC}$: Monte Carlo scale factor. The signal Monte Carlo was generated with statistics at a level of 200 times the standard solar model predicted flux for ^{8}B neutrinos and 2000 times for *hep* neutrinos.

$N_{\rm D}$: Number of target deuterons. As the CC and NC reactions proceed through reaction on deuterium it is important that the number of deuterons simulated in the detector is correct. Calculations indicated that a slight adjustment was required over the number used in the Monte Carlo.

$N_{\rm e}$: Number of target electrons. A small correction was required to bring the number of electrons in the Monte Carlo in line with the latest calculations. This only affected the ES reaction, which proceeds through reactions on electrons.

$S_{\rm ac}$: Data-cleaning correction. Although the criteria used to remove the unwanted instrumental backgrounds have been finely tuned [66], there is a finite chance that true neutrino events could be mis-identified as background events and removed from the data set. The probability that good neutrino events are removed in this way is defined as *sacrifice* and was studied using various calibration sources [100]. These studies were also carried out as a function of energy. Figure I.1 shows the energy dependence of sacrifice in the salt data set and table I.2 contains the differential sacrifice numbers, $S_{\rm ac}(T_{eff})$, for each signal type. The effect of sacrifice was included in the PDFs used in the signal extraction process as detailed in section 6.1.

$R_{\rm O,Na,Cl}$: Reactions on ^{18}O, Na, and Cl. The CC reaction can also proceed on ^{18}O, Na, and Cl nuclei in the D_2O volume. These additional reactions were not included in simulations but serve to slightly increase the CC flux.

$g_{\rm A}$ and $L_{1,\rm A}$: These are both corrections to bring the Monte Carlo in line with the latest cross-sections from effective field theory calculations [101]. $g_{\rm A}$ is the weak axial coupling and $L_{1,\rm A}$ is the two body axial exchange-current counter-term.

$MC_{\rm err}$: Aborted simulation events. A small decrease in the number of events simulated arises due to errors tracking events through the detector geometry, and errors in the data acquisition processor. The energy dependent factor required to account for the number of events lost is detailed in appendix F and was included in the signal extraction process.

$\frac{t'}{t}$: Livetime correction. This accounts for the loss in livetime due to the muon follower cut (described in appendix C).

I.2 Predicted Numbers of Events

Predictions for the expected number of CC and ES events (E(CC) and E(ES)) were obtained from the number of Monte Carlo simulated events passing event selection criteria, N(CC) and N(ES):

$$E^{\rm X}({\rm CC}) = \left(\frac{N({\rm CC})^{\rm X}_{\rm MC} R^{\rm X}_{\rm model}}{\alpha^{\rm X}_{\rm MC}} \right) N_{\rm D} S_{ac} g_{\rm A} L_{1,\rm A} R_{\rm O,Na,Cl} MC_{\rm corr} \frac{t'}{t} \tag{I.3}$$

$$E^{\rm X}({\rm ES}) = \left(\frac{N({\rm ES})^{\rm X}_{\rm MC} R^{\rm X}_{\rm model}}{\alpha^{\rm X}_{\rm MC}} + \right) N_{\rm e} S_{ac} MC_{\rm corr} \frac{t'}{t} \tag{I.4}$$

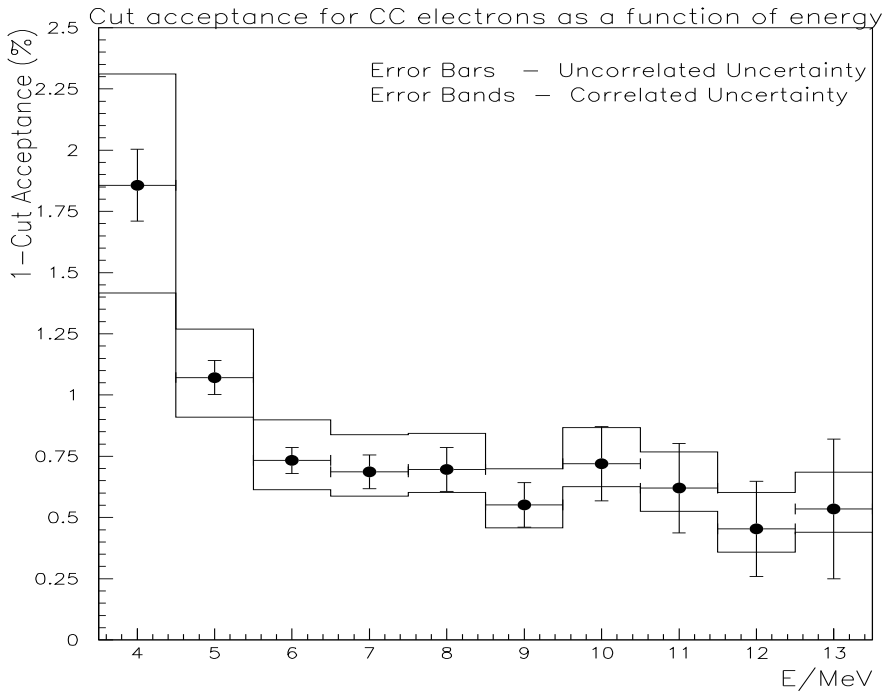


Figure I.1: Differential total sacrifice for CC electrons. (note that the horizontal axis shows E = T_{eff}). Both correlated and uncorrelated errors are shown in this plot taken from Ref. [100].

T_{eff} (MeV)		Sacrifice		
Minimum	Maximum	CC	ES	NC
3.5	4.5	$0.01763^{+0.00479}_{-0.00465}$	$0.02979^{+0.00930}_{-0.00923}$	$0.00682^{+0.00154}_{-0.00100}$
4.5	5.5	$0.00978^{+0.00214}_{-0.00180}$	$0.02033^{+0.00567}_{-0.00555}$	$0.00682^{+0.00154}_{-0.00100}$
5.5	6.5	$0.00638^{+0.00177}_{-0.00136}$	$0.01347^{+0.00385}_{-0.00366}$	$0.00682^{+0.00154}_{-0.00100}$
6.5	7.5	$0.00592^{+0.00170}_{-0.00125}$	$0.00970^{+0.00248}_{-0.00220}$	$0.00682^{+0.00154}_{-0.00100}$
7.5	8.5	$0.00601^{+0.00177}_{-0.00135}$	$0.00843^{+0.00210}_{-0.00176}$	$0.00682^{+0.00154}_{-0.00100}$
8.5	9.5	$0.00456^{+0.00177}_{-0.00136}$	$0.00563^{+0.00186}_{-0.00147}$	$0.00682^{+0.00154}_{-0.00100}$
9.5	10.5	$0.00624^{+0.00215}_{-0.00182}$	$0.00646^{+0.00216}_{-0.00183}$	$0.00682^{+0.00154}_{-0.00100}$
10.5	11.5	$0.00524^{+0.00238}_{-0.00208}$	$0.00564^{+0.00240}_{-0.00211}$	$0.00682^{+0.00154}_{-0.00100}$
11.5	12.5	$0.00359^{+0.00238}_{-0.00208}$	$0.00377^{+0.00240}_{-0.00211}$	$0.00682^{+0.00154}_{-0.00100}$
12.5	13.5	$0.00440^{+0.00248}_{-0.00220}$	$0.00498^{+0.00250}_{-0.00222}$	$0.00682^{+0.00154}_{-0.00100}$
13.5	20.0	$0.00000^{+0.00324}_{-0.00303}$	$0.00000^{+0.00330}_{-0.00309}$	$0.00682^{+0.00154}_{-0.00100}$

Table I.2: Differential sacrifice values for the CC, ES and NC signals in terms of observable kinetic energy. The errors are combined correlated and uncorrelated uncertainties.

where X is the type of neutrino produced in the Sun, and N(CC) and N(ES) include the energy-dependent correction factors applied in the MXF code. ^{8}B and *hep*-type neutrinos were treated separately in these calculations because their predicted fluxes differ significantly.

The number of NC events predicted, E(NC), was obtained from:

$$E(\mathrm{NC}) = R_{\mathrm{NC}}\epsilon_{\mathrm{n}}t'\omega_{\mathrm{NC}}^{\mathrm{RC}}N_{\mathrm{D}}S_{ac}g_{\mathrm{A}}L_{1,\mathrm{A}}R_{\mathrm{model}}^{\mathrm{B}} \tag{I.5}$$

where

$R_{NC} = \mathbf{13.107738}$ is the total predicted rate of neutrons per day assuming the SSM ^{8}B and *hep* fluxes [31],

$\epsilon_{\mathrm{n}} = \mathbf{0.398 \pm 0.005^{+0.012}_{-0.010}}$ is the neutron detection efficiency inside all data cuts [31], and

$t' = \mathbf{391.43\ days}$ is the corrected livetime of the salt data set.

As the model correction factor is higher for the ^{8}B flux than the *hep* flux, this calculation will slightly over-estimate the predicted number of NC events. However, as the *hep* flux is so much smaller than the ^{8}B flux ($\approx 0.14\%$) this will be a small effect.

I.3 Flux Normalisation Factors

The flux of neutrinos causing a given reaction type in SNO was calculated by multiplying the number of events detected by a normalisation factor. The normalisation required was

$$\mathcal{N}(Z) = \frac{\Phi^{\mathrm{B}}_{\mathrm{SSM}} + \Phi^{\mathrm{hep}}_{\mathrm{SSM}}}{E^{\mathrm{B}}(Z) + E^{\mathrm{hep}}(Z)}, \tag{I.6}$$

where Z is either CC or ES, the Φ_{SSM} factors are the BP2004 model predictions and $E(Z)$ are the expected numbers of events as explained in the previous section.

Since the ^{8}B and *hep* values were combined in R_{NC} the normalisation factor for the NC flux was:

$$\mathcal{N}(NC) = \frac{\Phi^{\mathrm{B}}_{\mathrm{SSM}} + \Phi^{\mathrm{hep}}_{\mathrm{SSM}}}{E(\mathrm{NC})}. \tag{I.7}$$

Appendix J

A Likelihood Fit Without Neutron Energy Information

In order to check the validity of fixing the neutron energy spectrum in the fit, a completely unconstrained fit was carried out in which the neutron flux was varied separately in each 0.5 MeV bin. The likelihood for each event was obtained from:

$$\mathcal{L} = \sum_{i=\mathrm{CC,ES,NC,EN}} \left(\sum_{k=1}^{N_s} \left(n_{i_k} . P(i_k/T_{eff}, \beta_{14}, R^3) . P(i_k/\cos\theta_\odot/R^3) \right) \right) \qquad \text{(J.1)}$$

A separate fit was carried out for data falling into 12 different bins in T_{eff}. The first 11 bins ranged from 5.5–11.0 MeV and were 0.5 MeV wide, whilst the last bin ran from 11.0–20.0 MeV, due to the low statistics of neutrons in this energy region. Internal γ events were omitted from this fit as the contribution of these events to each separate fit is small ($\lesssim 2$ events). Apart from this the extraction was run in the same way as prescribed for the regular unconstrained fit presented in chapter 6.

The fit results are plotted in figure J.1 and the total fitted numbers of events for each event-class is given in table J.1. The uncertainties given in this table are the quadrature sum of the statistical uncertainties in each bin. The results are in good agreement with those obtained from the fit including the neutron energy information (also given in table J.1 for comparison) indicating that there are no large deviations from the expected neutron energy distribution. The noticeable increase in the magnitude of the statistical uncertainties for the fitted CC and ES signals is indicative of the effect of the neutron energy information with-held from the fit.

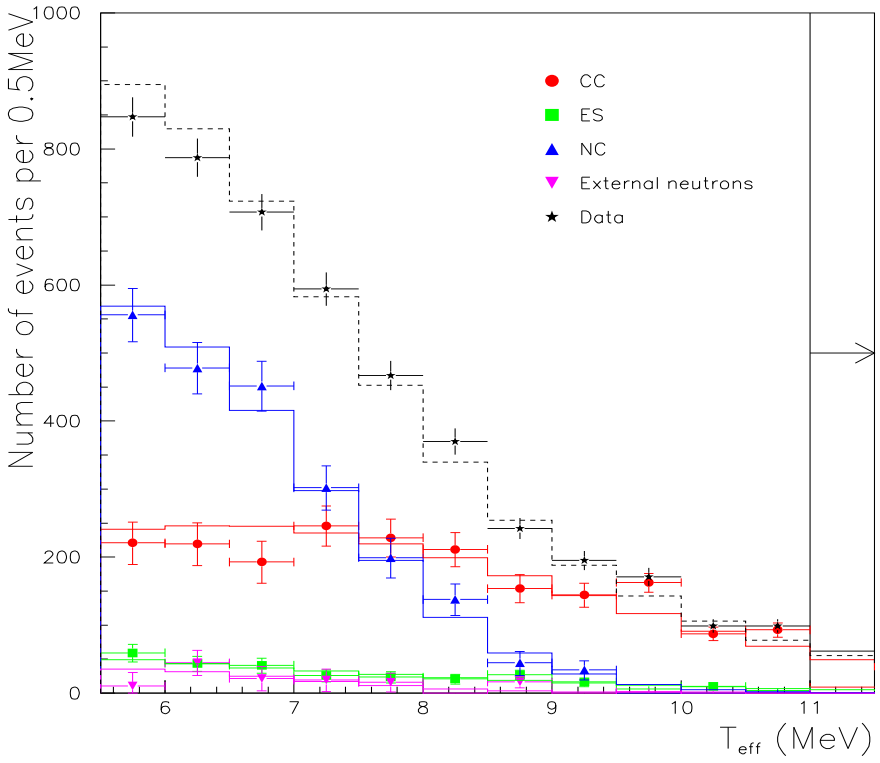


Figure J.1: Fitted events for each of the signals in 0.5 MeV bins when the neutron energy information was not used to constrain the fit. Only statistical errors are shown.

Signal	Number of Events Fitted	
	With neutron energy information	No Neutron energy information
CC	$2094.6\ ^{+79.7}_{-78.9}$	$2178.71^{+81.55}_{-79.49}$
ES	$284.0\ ^{+26.9}_{-27.3}$	$288.29^{+29.45}_{-26.50}$
NC	$2088.4\ ^{+91.4}_{-90.8}$	$2128.32^{+86.50}_{-83.07}$
EN	$133.5\ ^{+42.8}_{-41.8}$	$151.69^{+42.36}_{-40.38}$

Table J.1: The total fitted events for each signal type for the regular energy-unconstrained fit presented in chapter 6, and a fit in which neutron energy information is not included. Only statistical errors are given here.

Bibliography

[1] W. Pauli. In a letter to the Physical Institute of the Federal Institute of Technology in Zürich, *Physics Today* (1978).

[2] F. Reines & C.L. Cowan. *Nature*, **178** (1956) 446.

[3] B.T. Cleveland *et al. Astrophys. J.*, **496** (1998) 505–526.

[4] B. Pontecorvo. *Soviet Physics JETP*, **26**, 5 (1968) 984.

[5] V. Gribov & B. Pontecorvo. *Phys. Lett. B*, **28**, 7 (1969) 493.

[6] F. Halzen & A. Martin. "Quarks & Leptons: An Introductory Course in Modern Particle Physics". John Wiley & Sons, Inc. (1984).

[7] K. Hagiwara *et al. Phys. Rev. D*, **66** (2002) 010001.

[8] J. Bonn *et al. Nuclear Physics B Proc. Suppl.*, **91** (2001) 273–279.

[9] V.M. Lobashev *et al. Nuclear Physics B Proc. Suppl.*, **91** (2001) 280–286.

[10] K. Assamagan *et al. Phys. Rev. D*, **53** (1996) 6065.

[11] R. Barate *et al. Eur. Phys. J. C*, **2** (1998) 395. ALEPH-coll.

[12] S.F. King. *Rept. Prog. Phys.*, **67** (2004) 107–158.

[13] O. Cremonesi. *Nuclear Physics B Proc. Suppl.*, **118** (2003) 287–296.

[14] D.N. Spergel *et al. Astrophys.J.Suppl.*, **148**, 2003 (2003) 175.

[15] C. Athanassopoulos *et al. Phys. Rev. Lett.*, **81** (1998) 1774.

[16] R. Tayloe for the MiniBooNE Collaboration. *Nuclear Physics B Proc. Suppl.*, **118** (2003) 157–163.

[17] Y. Fukuda *et al* Super-Kamiokande collaboration. *Phys. Rev. Lett.*, **81** (1998) 1562–1567.

[18] M. Altmann *et al. Phys. Lett. B*, **490** (2000) 16.

[19] V.N. Gavrin for the SAGE collaboration. *Nuclear Physics B Proc. Suppl.*, **118** (2003) 39–46.

[20] Y. Fukuda *et al. Phys. Rev. Lett.*, **77** (1996) 1683–1686.

[21] J.N. Bahcall & S. Basu. *Phys. Rev. Lett.*, **92** (2004) 121301.

[22] J.N. Bahcall. "Neutrino Astrophysics". Cambridge University Press (1989).

[23] J.N. Bahcall, M.N. Pinsonneault & S. Basu. *Astrophys. J.*, **555**.

[24] A.R. Junghans *et al. Phys. Rev. C*, **68** (2003) 065803.

[25] J.N. Bahcall & C. Pe na Garay. *New Journal of Physics*. ArXiv:hep-ph/0404061.

[26] L. Wolfenstein. *Phys. Rev. D*, **17** (1978) 2369.

[27] S.P. Mikheev & A.Y. Smirnov. *Soviet Journal of Nuclear Physics*, **46(2)** (1985) 913.

[28] Q.R. Ahmed *et al. Phys. Rev. Lett.*, **87** (2001) 071301.

[29] C.E. Ortiz *et al. Phys. Rev. Lett.*, **85** (2000) 2909.

[30] The SNO collaboration. *Phys. Rev. Lett.*, **89** (2002) 011301.

[31] The SNO collaboration. *Phys. Rev. Lett.*, **92** (2004) 181301.

[32] K. Eguchi *et al. Phys. Rev. Lett.*, **90** (2003) 021802.

[33] P.C. de Holanda & A.Y. Smirnov. *JCAP*, **0302** (2003) 001.

[34] S.T. Petcov & S. Toshev. *Phys. Lett. B*, **187** (1987) 120.

[35] M. Apollonio *et al. Eur. Phys. J.*, **C27** (2003) 331.

[36] J.N. Bahcall, M.C. Gonzalez-Garcia & C. Pe na Garay. *JHEP*, **0302** (2003) 009.

[37] J. Schechter & J.W.F. Valle. *Phys. Rev. D*, **24** (1981) 1883.

[38] E.Kh. Ahkmedov. *Phys. Lett. B*, **213** (1988) 64.

[39] J.W. Valle. *Phys. Lett. B*, **199** (1987) 432.

[40] K. Zuber. "Neutrino Physics". IoP publishing (2004).

[41] The SNO collaboration. *Phys. Rev. Lett.*, **89** (2002) 011302.

[42] M.B. Smy. *Nucl. Phys. B Proc. Suppl.*, **118**, 2003.

[43] H.H. Chen. *Phys. Rev. Lett.*, **55** (1985) 1534–1536.

[44] SNO Collaboration. *Nucl. Instrum. Meth.*, **A449** (2000) 172–207.

[45] T.C. Andersen *et al. Nucl. Instrum. Meth.*, **A501** (2003) 399–417.

[46] B. Hyland. "Final Results for the MnOx D_2O Ra Assays for the Pure D_2O Phase". Technical report, SNO (2002).

[47] I. Lawson. "Results of the MnOx D_2O Assays During the Salt Phase of the SNO Experiment". Technical report, SNO (2003).

[48] T.C. Andersen *et al. Nucl. Instrum. Meth.*, **A501** (2003) 386–398.

[49] SNO Low Energy Background Analysis Group. "Low Energy Background in the Salt Phase of the SNO Experiment." Technical report, SNO (2004).

[50] M.R. Dragowsky *et al. Nucl. Instrum. Meth.*, **A481** (2002) 284–296.

[51] N. Tagg *et al. Nucl. Instrum. Meth.*, **A489** (2002) 92–102.

[52] A.W.P. Poon *et al. Nucl. Instrum. Meth.*, **A452** (2000) 15–129.

[53] CERN Program Library Long Write Up. "Zebra Q100/Q101" (1995).

[54] M.G. Boulay. "Direct Evidence for Weak Flavour Mixing with the Sudbury Neutrino Observatory". Ph.D. thesis, Queen's University (2001).

[55] W.R. Nelson. "The EGS4 Code System". Technical report, SLAC (1985).

[56] M.D. Lay. "Creation and Detection of Čerenkov light in the Sudbury Neutrino Observatory." Ph.D. thesis, University of Oxford (1994).

[57] J.A. Dunmore. "The Separation of CC and NC Events in the Sudbury Neutrino Observatory". Ph.D. thesis, University of Oxford (2004).

[58] M. Thorman. Ph.D. thesis, University of Oxford. Not yet submitted.

[59] "A Monte Carlo N-Particle Transport Code System, *version 4C*". Technical report, Los Alamos National Laboratory (1993).
`http://www-xdiv.lanl.gov/XCI/PROJECTS/MCNP/`.

[60] The SNO collaboration. "The SNOMAN Companion", 4.01 edition.

[61] CERN Program Library Long Write Up. "HEPDB - Database Management Package, Q180 *version 1.9*" (1995).

[62] K. Graham. "Salt Electron Energy Response Update Note". Technical report, SNO (2003).

[63] P. Vogel & J. Beacom. *Phys. Rev. D*, **60** (1999) 053003.

[64] V.L. Rusu. "Measurement of the total ^{8}B solar neutrino flux at the Sudbury Neutrino Observatory". Ph.D. thesis, University of Pennsylvania (2003).

[65] M. Boulay *et al.* "Other Background Sources for the Extended Salt Data". Technical report, SNO (2004).

[66] S. Brice *et al.* "Sacrifice and Contamination in the SNO Data Cleaning Cuts." Technical report (2000). SNO internal analysis note.

[67] International Commission on Radiation Units & Measurements. *ICRU Report*, **37**.

[68] R.J. Boardman. "The Detection of Čerenkov Radiation from Neutrino Interactions". Ph.D. thesis, Oxford University (1992).

[69] J.A. Dunmore (2001). Private communication.

[70] M.G. Bowler (2001). Private communication.

[71] M.G. Bowler & M.D. Lay. *Nucl. Instrum. Meth. sec A*, **378** (1996) 468–471.

[72] M.S. Neubauer. "Evidence for electron neutrino flavor change through measurement of the ^{8}B solar neutrino flux at the Sudbury Neutrino Observatory". Ph.D. thesis, University of Pennsylvania (2001).

[73] A. Hallin. "Cross Talk IV". Technical report, SNO (2001). `http://www.sno.phy.queensu.ca/~hallin/private/nu/CrossTalkIV.html`.

[74] A. Hamer. "Estimates of Systematic Uncertainties in Energy for an Nhit Based Analysis. *version II*". Technical report, SNO (2002).

[75] J.E. Perry & Jr S.J. Bame. *Phys. Rev.*, **99** (1955) 1368.

[76] X. Dai. "A Chemical Explanation of Heavy Water Attenuation". Technical report, SNO (2002).

[77] E. Hecht. "Optics". Addison-Wesley, third edition (1998).

[78] M. Pickel. "Rayleigh scattering in heavy and light water". Technical report, Queens University (2000).

[79] R.C. Smith & K.S. Baker. *Applied Optics*, **20(2)** (1981) 177–184.

[80] H. Buiteveld *et al. Ocean Optics XII*, **2258** (1994) 174–183. Society of photo-optical instrumentation engineers.

[81] R.J. Ford. "Calibration of SNO for the Detection of ^{8}B Neutrinos". Ph.D. thesis, Queens University (1998).

[82] G. Cowan. "Statistical Data Analysis". Clarendon Press, Oxford (1998).

[83] F. James. "CERN Program Library Long Writeup D506". Technical report, CERN (1998).

[84] J. Wilson. "MXF CODE documentation" (2003). `http://www-pnp.physics.ox.ac.uk/~wilson/work/mxf_documentation.html`.

[85] W.H. Press *et al.* "Numerical Recipes in Fortran 77". Cambridge University Press, second edition (1992).

[86] P.R. Bevington & D.K. Robinson. "Data Reduction and Error Analysis for the physical sciences". McGraw-Hill, third edition (2003).

[87] K. Schaffer & S. Oser. "Livetime numbers for the final salt data set". Technical report, SNO (2003).

[88] K. Graham. "Energy Systematics for the NSP". Technical report, SNO (2004).

[89] G. McGregor. "The Measurement of the Neutral Current Interaction at the Sudbury Neutrino Observatory". Ph.D. thesis, University of Oxford (2002).

[90] K. Graham. "β_{14} Response from NSP". Technical report, SNO (2004).

[91] A. Hallin. "^{8}Li Source and β_{14} as a Function of Energy". Technical report, SNO (2004).

[92] J. Wilson. "A Quick Note on Energy Dependent β_{14} Systematic Uncertainty". Technical report, SNO (2004).

[93] K. Graham. "NSP Angular Response Update". Technical report, SNO (2004).

[94] M. Boulay *et al.* "Checks on Reconstruction Using the AV Location". Technical report, SNO (2003).

[95] K. Miknaitis & A. Hallin. "Vertex Shift Studies Using ^{16}N in Salt". Technical report, SNO (2003).

[96] M. Kos. "Neutron Detection Efficiency for the Complete Salt Data Set". Technical report, SNO (2004).

[97] J.R. Klein. "Update to Monte Carlo Estimate of the Čerenkov Background Produced by Internal D_2O Radioactivity for the Next Salt Paper". Technical report, SNO (2004).

[98] J.R. Klein. "Update to Bi Internal Tail Using Spike Data". Technical report, SNO (2004).

[99] M. Dunford. "Measurement of External Čerenkov Events for the Full Salt Phase". Technical report, SNO (2004).

[100] M. Dunford *et al.* "Updates to Data Cleaning for the Salt Phase - *Version 1.2*". Technical report, SNO (2004).

[101] M. Butler, J.W. Chen & X. Kong. *Phys. Rev. C*, **63** (2001) 035501.

[102] S. Nakumura *et al. Nucl. Phys.*, **A707** (2002) 561.

[103] A. Kurylov, M.J. Ramsey-Musolf & P. Vogel. *Phys. Rev.*, **C65** (2002) 05501.

[104] SNO Day-Night Working Group. "Signal Extraction and Results for the Salt Day-Night Analysis". Technical report, SNO (2004).

[105] The KamLAND collaboration. *Submitted to Phys. Rev. Lett.*. arXiv:hep-ex/0406035.

[106] G. Tesic. "NSP physics interpretation tables and plots". Technical report, SNO (2004).
`http://www.physics.carleton.ca/~gtesic/qphys/qphys.html`.

[107] H. Robertson. "Long D_2O Paper: Target Properties". Technical report, SNO (2003).

[108] P.C. de Holanda & A.Yu. Smirnov. *Phys. Rev. D*, **69** (2004) 113002.

[109] M. Maltoni *et al. Phys. Rev. D*, **68** (2003) 113010.

[110] E. Kh. Akhmedov, M.A. Tórtola & J.W.F. Valle. *JHEP*, **405** (2004) 57.

[111] G. Alimonti *et al.* Borexino Collaboration. *Astropart.Phys.*, **16** (2002) 205–234.

[112] The K2K Collaboration. *Phys. Rev. Lett.*, **90** (2003) 041801.

[113] The MINOS collaboration. "MINOS technical design report". Technical report (1998).
`http://www.hep.anl.gov/ndk/hypertext/minos_tdr.html`.

[114] K. Anderson *et al.* "White paper report on using nuclear reactors to search for a value of θ_{13}". Technical report (2004).
`http://www.hep.anl.gov/minos/reactor13/reactor13.pdf`.

[115] S. Eidelman *et al. Phys. Lett. B*, **592** (2004) 1.

[116] X. Chen. "Monte Carlo and Analysis Techniques for the Sudbury Neutrino Observatory". Ph.D. thesis, Oxford University (1996).

[117] P. Harvey. "Rootification livetime tool." `http://www.sno.phy.queensu.ca/~phil/private/cgi-bin/rootify.cgi`.

[118] N.K. McCauley. "Producing a Background Free Data Set for Measurement of the Charge Current Flux and Day-Night Asymmetry at the Sudbury Neutrino Observatory". Ph.D. thesis, University of Oxford (2001).

[119] W. Heintzelman. "Use of Angular Correlation Functions to separate CC and NC events in Salt." Technical report (2002). SNO-STR-2002-004.

[120] M. Dunford (2004). Private communication.

[121] A. Hallin. "Cross Talk IV". Technical report, SNO (2001). `http://www.sno.phy.queensu.ca/~hallin/private/nu/CrossTalkIV.html`.

[122] J. Hall (2004). Private communication.

[123] N. Tagg. "The ^{8}Li Calibration Source and Through-going Muon Analysis in the Sudbury Neutrino Observatory". Ph.D. thesis, University of Guelph (2001).

[124] J. Wilson. "An attempt to study the Energy Nonlinearity Systematic with ^{8}Li - *Version III*". Technical report, SNO (2004).

[125] K. Graham. "Simple ^{8}Li Selection". Technical report, SNO (2004).

[126] M. Boulay. "An Energy Calibration of SNO - The RSP Processor *DRAFT*". Technical report, SNO (2002).

[127] J. Wilson. "Studies into Energy Dependence of Vertex Shift". Technical report, SNO (2004).

[128] J.N. Bahcall, M.H. Pinsonneault & S. Basu. *Astrophys.J.*, **555** (2001) 990–1012.

[129] Y. Takeuchi. "Cross section correction factor for salt ES data". Technical report, SNO. `http://owl.phy.queensu.ca/~takeuchi/private/esrd/`.