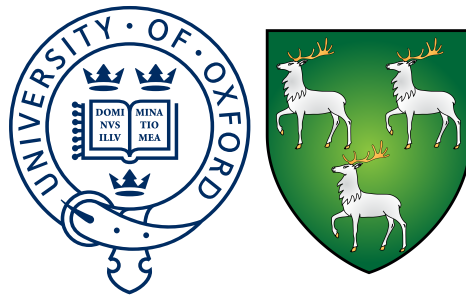# A Measurement of Scattering Characteristics of the Detection Medium in the SNO+ Detector

Esther Turner

Jesus College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2021

Abstract

The SNO+ experiment is a general purpose, low energy neutrino experiment which has repurposed the nobel prize winning SNO detector. The heavy water target is replaced with three different targets in phases: first, ultra-pure water (UPW); second, a liquid scintillator; and finally, the liquid scintillator loaded with ^{nat}Te. This final phase is focused on the primary physics goal of SNO+: to search for neutrinoless double beta decay in ^{130}Te. Reaching a competitive sensitivity to this process requires a full understanding of the detector response to various particle interactions. Therefore, a comprehensive calibration campaign is conducted in each of the three phases, both for the physics analyses occurring in that phase and also to build up understanding of the detector response through the more well understood targets.

This thesis concerns the calibration of the scattering properties of the detector using a detector system called SMELLIE - the Scattering Module of the Embedded LED/Laser Light Injection Entity. SMELLIE enables an in-situ measurement of the scattering characteristics of the target medium by firing pulsed lasers with a wavelength range of 375 nm - 695 nm across the detector. While the majority of the system had been installed prior to this work, SMELLIE had not yet been commissioned. Therefore this thesis covers the initial commissioning of SMELLIE, including a method for firing a specified shot energy at any available wavelength, investigating unusual beam features and understanding the timing spectrum of detector PMT hits during a SMELLIE run. Crucially, a method is also developed for accurately modelling the angular emission profiles of the beams in the detector Monte Carlo simulation.

The final goal of this work is the analysis of the scattering characteristics of the UPW target: these were found to be consistent with Rayleigh scattering, as expected. The Rayleigh scattering length has been measured

at range of wavelengths (375, 435, 495 and 555 nm), has since been implemented in the detector simulation software as a divisive scaling factor of $1.28 \pm 0.05(\text{stat.}) \pm 0.14(\text{sys.})$ applied to the default Rayleigh scattering length (which was extrapolated from SNO values) and used in the global calibration campaign for the UPW phase. This analysis also resulted in scattering measurements of the water external to the target volume and was crucial to determining the optical model of the target vessel, both of which will be used for future phases.

Acknowledgements

Firstly, many thanks go to my supervisor, Armin Reichold. This thesis took longer than either of us expected and you have continued to care about my wellbeing and provide very thorough feedback throughout.

Thank you also to my original supervisor, Steve Biller, without whose enthusiasm I would never have ended up working on SNO+. And to Jeff Tseng - I really enjoyed the chance to work with you and your summer students.

I believe I had one of the best PhD offices that I could possibly have asked for: Krish Majumdar, Chris Jones, Jack Dunger, Jeff Lidgard, Luca Cavalli, Tereza Kroupa, Iwan Morton-Blake, Josh Wang and Josie Paton. Thanks to all of you for making my time in the office a time I look back on very fondly. Particular thanks to Jack Dunger, for your continued interest in my work from the moment I joined to you leaving the country; Jeff Lidgard, my SMELLIE buddy, for all those underground shifts in which you taught me more about optical circuitry than I'd ever expected to learn; Iwan Morton-Blake, for becoming such a surprisingly great friend; and Tereza Kroupa - I am so fortunate that you chose to join our group for your DPhil as I cannot imagine having done mine without you. Your company in the office, writing sessions and sleepovers mean more than I can say. Not to mention the weekly chats, laughs and beers along with Iwan.

Not in the office, but equally worthy of thanks is Ed Leming. You always said I'd get to this point one day.

This thesis wouldn't have been possible without the rest of the SNO+ collaboration. My thanks to you all, particularly the Optics group. Also, thanks to those who made trips to Sudbury as fun as possible - Billy Liggins in particular.

Elsewhere in the physics department, many thanks go to Kim Proudfoot and Sue Geddes, who have always been so approachable, helpful and excellent at solving problems. Also to Theresa Fruth - you've been wonderful over the last seven years.

Of course, many thanks to my parents, who have always supported me through this and without whom I probably wouldn't have ended up doing a DPhil in the first place. And to the rest of my family as well - you have all put up with a lot. Sorry about

those family holidays when I worked rather than spending time with you all. Also, Andy and Heather, whose friendship and DPhil survival kits have been invaluable.

And finally and most importantly, James. I could not have done this without you by my side cheering me on all the way. You had to listen to so much more complaining than your share, put up with very late work nights, my not being paid for six months and, of course, my trips to Canada. But all through it, you believed in me and I owe you. Thank you so much for everything.

Contents

List of Figures

Chapter 1

Neutrino Physics

1.1   A History of Neutrinos

Neutrinos were first proposed by W. Pauli in 1930 [1] as an attempt to maintain energy conservation in β decays. Rather than a single, sharp energy peak, as expected from a two body decay, a continuous electron-energy spectrum was observed [2]. Pauli suggested that a low mass, electrically neutral, spin $\frac{1}{2}$, weakly interacting particle could be created during β decay and carry away the remaining energy without being observed.

E. Fermi then developed a theory of weak interactions [3], in which β decay is a 4-fermion interaction including the decaying neutron or nucleus and the three outgoing fermions (electron, proton or daughter nucleus and the neutrino). Following this, H. Bethe and R. Peierls calculated the cross section for $\nu_e + n \rightarrow e^- + p$ to be 10^{-44} cm^2 at a neutrino energy of a few MeV [4]. This extremely small cross section led them to believe that neutrinos would never be observed.

In fact, the neutrino was detected 26 years after it was postulated. C. L. Cowan and F. Reines detected $\overline{\nu}_e$ in 1956 [5, 6] by using inverse β decay $(\overline{\nu}_e + p \rightarrow e^+ + n)$ interactions on a water target. In order to overcome the extremely small cross-section, they used a nuclear reactor at Savannah River, North Carolina, as an extremely high flux $\overline{\nu}_e$ source.

After the experimental discovery of $\overline{\nu}_e$, ν_μ were discovered at Brookhaven in 1962 [7] and ν_τ were observed by the DONUT collaboration in 2000 [8].

The Z boson resonance observation at LEP was used to measure the number of light neutrino flavours which interact via the weak force [9]. This was found to be consistent with three (i.e. the ν_e, ν_μ and ν_τ).

Having briefly discussed the discovery of neutrinos, the discovery of the properties of neutrinos and the weak interaction will be covered next.

In 1956, C. N. Lee and T. D. Yang proposed that parity may be violated in weak interactions [10]. C. S. Wu then designed an experiment to test this proposal [11]. She and her collaborators used ^{60}Co atoms (which undergo β decay) with their spins aligned in a magnetic field. The direction of the outgoing electron from the β decay was observed. Rather than electrons being emitted isotropically, as would be expected from parity conservation, electrons were preferentially emitted in the direction of the ^{60}Co spin vector. This observation led to the development of the V-A theory of weak interactions [12–14], which predicts maximal violation of parity.

Using electron capture on ^{152m}Eu in 1957, M. Goldhaber and his collaborators measured the helicity of neutrinos [15]. ^{152m}Eu decays to ν_e and ^{152m}Sn*, which then decays to its ground state by emitting a γ. The polarisation of this γ was measured and showed that all the neutrinos produced had left-handed helicity.

In the Standard Model, the charged weak interaction only couples to left-handed chiral particles (or right-handed anti-particles) [16–18]. In addition, in the massless limit, helicity becomes identical to chirality. Therefore, always observing neutrinos with left-handed helicity implies that neutrinos are massless.

Indeed, neutrinos were considered to be massless until the discovery of neutrino oscillations in 2001, discussed in the following section, which are the only terrestrial beyond the Standard Model (BSM) physics which has been observed to date.

1.2 Neutrino Oscillations

The discovery of neutrino mass was a result of two observations: a deficit of observed solar ν_e observed compared to predictions from the standard solar models [19] and a deficit of observed atmospheric ν_μ [20] when compared to expectations. These problems were known as the solar neutrino problem and the atmospheric neutrino anomaly, respectively.

Both of these anomalies could be explained through the phenomenon of neutrino-flavour mixing, analogous to the mixing which occurs in kaon oscillations. Flavour mixing can be explained by assuming that the weak flavour eigenstates are not the same as the mass eigenstates and that neutrinos therefore change flavour as they propagate. Clearly, this requires that neutrinos have a non-zero mass (specifically, that two of the three mass eigenstates are non-zero and not identical).

Each mass eigenstate is a linear combination of the three flavour eigenstates [21]. This is described by

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{1.1}$$

where

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix}. \tag{1.2}$$

In Eq. 1.1, ν_i ($i = 1, 2, 3$) denotes the mass eigenstates and ν_α ($\alpha = e, \mu, \tau$) denotes the flavour eigenstates. U is the lepton mixing matrix, also known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [21].

As each component of U is a complex number, there could, at most, be 18 real parameters in this matrix. However, U is unitary as it describes the decomposition of probabilities [21]. Therefore, the number of parameters is reduced to nine and by applying unitary diagonal transforms all but one complex phase can be cancelled [21]. This leads to the conventional parameterisation of the PMNS matrix:

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{CP}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{CP}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{CP}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{CP}} & c_{23}c_{13} \end{pmatrix}. \tag{1.3}$$

In Eq. 1.3, $c_{ij} = \cos(\theta_{ij})$ and similarly $s_{ij} = \sin(\theta_{ij})$. θ_{ij} are the mixing angles and δ_{CP} is the charge-parity (CP) violating phase.

In practice, lepton mixing means that when a $W^\pm$ couples to a flavour eigenstate of both a charged and neutral lepton, the neutral lepton is actually the superposition of the three neutrino mass eigenstates [21]. These mass eigenstates, provided they are not identical but carry the same momentum, propagate at different speeds. Therefore, phase differences arise between the different mass eigenstates in the wave-function of the flavour eigenstate. Consequently, when the neutrino is detected, the probabilities of each flavour eigenstate appearing have changed since its creation. This means that the neutrino detected can be a different flavour to the neutrino produced - this is neutrino oscillation.

Neutrino oscillations are described by

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta} = \delta_{\alpha\beta} &- 4 \sum_{i>j} \mathrm{Re}(U_{\alpha i}^* U_{\beta i}^* U_{\alpha j} U_{\beta j}^*) \sin^2 \left(\frac{\Delta_{ij}^2 L}{4E} \right) \\ &+ 2 \sum_{i>j} \mathrm{Im}(U_{\alpha i}^* U_{\beta i}^* U_{\alpha j} U_{\beta j}^*) \sin \left(\frac{\Delta_{ij}^2 L}{2E} \right), \end{aligned} \tag{1.4}$$

where $P_{\nu_\alpha \to \nu_\beta}$ is the probability that a neutrino of energy E produced in flavour eigenstate α has flavour β after having travelled distance L [21] (known as the baseline in an oscillation experiment). Δ_{ij}^2 is the difference in the squared masses between mass eigenstate i and j.

It can be seen from Eq. 1.4 that neutrino oscillation experiments can be used to measure components of the PMNS matrix and the squared mass differences. However, neutrino oscillation experiments, crucially, cannot be used to measure the absolute masses of the neutrinos.

In 2015, T. Kajita and A. B. McDonald were awarded the Nobel Prize in Physics for the discovery of neutrino oscillations and hence the discovery of neutrino mass. These men represent the Super-Kamiokande and SNO experiments respectively.

Super-Kamiokande, a large water Cherenkov detector, measured the atmospheric neutrino flux. It found that the ν_e flux was consistent with expectations, whereas the ν_μ flux was lower than expected. However, this deficit was dependent on the direction of the neutrino entering the detector [22]. Having been created in the atmosphere, neutrinos that enter the detector from above have a much shorter baseline than neutrinos which enter from below (and have had to travel through the Earth). The baseline dependence observed in the flux indicated the disappearance of ν_μ neutrinos consistent with oscillations [22].

SNO, using heavy water as a target and hence being sensitive to neutral current interactions, was able to make a flavour independent measurement of the solar neutrino flux [23]. The measured total neutrino flux was consistent with the Standard Solar Model [23]. However, earlier experiments (Homestake, GALLEX, SAGE, and Kamiokande) had measured the ν_e flux to be approximately $\sim 30~\%$ of expectation [21]. The combination of these results is evidence of ν_e oscillating into the other neutrino flavours [23].

The elements of the PMNS matrix are measured through a variety of oscillation experiments, with different baselines and energies used to access different matrix elements. θ_{12} and Δ_{12}^2 are the primary parameters controlling 'solar' oscillations, i.e. neutrino oscillations with baselines that match solar neutrinos. θ_{23} and Δ_{23}^2 are the primary atmospheric oscillation parameters and θ_{13} and Δ_{13}^2 are the reactor oscillation parameters.

Using neutrino vacuum oscillations, it is only possible to measure the magnitude of the mass squared differences [21] and not its sign. However, this is not necessarily true when matter effects are taken into account in the oscillation probabilities. The Mikheyev-Smirnov-Wolfenstein (MSW) effect is one such matter effect wherein
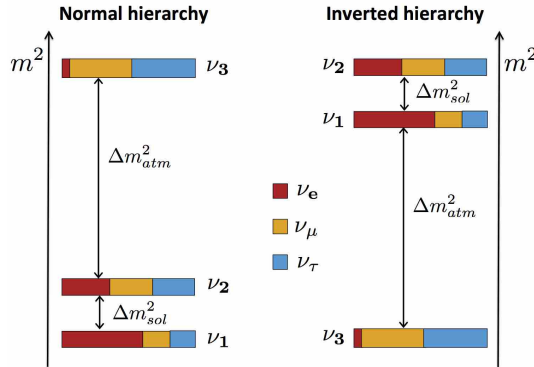
Figure 1.1: The two possible neutrino mass hierarchies: normal and inverted. Each mass eigenstate is coloured by its fraction of each flavour eigenstate. Figure from [27].

the electron density of the matter that the neutrinos propagate through causes the oscillation probability to depend on the sign of the mass squared difference. This is exploited by the solar oscillation experiments as the Sun is large and dense enough for the effect to be measurable [24].

The measurements from Super-Kamiokande, SNO, GALLEX, SAGE, Homestake and other oscillation experiments are combined into global neutrino oscillation fits, summarised in [25].

1.3 Unknowns in Neutrino Physics

As previously discussed, neutrino oscillations allow the measurement of the mixing angles, Δ_{12}^2, $|\Delta_{13}^2|$ and $|\Delta_{23}^2|$. However, there are still many unknowns in the field of neutrino physics. These include the value of δ_{CP}, which current experiments are starting to hint at [26], the octant of θ_{23} (i.e. whether it is greater, smaller or exactly 45°), the absolute masses of neutrinos, the mass hierarchy (i.e. whether Δm_{23}^2 is greater or smaller than zero) and whether neutrinos are Majorana particles. The latter three of these unknowns are the ones of interest for this work.

There are two possibilities for the mass hierarchy: the normal hierarchy and the inverted hierarchy. These two possible mass orderings are shown in Fig. 1.1. These are the only two possibilities due to the ability to measure the sign of Δ_{12}^2. This knowledge of the possible mass hierarchies and the squared mass differences, in turn, causes constraints on the possible absolute masses of neutrinos (which can be seen in Fig. 1.6).

There is an experimental effort to directly measure the absolute mass of the neutrino. This is traditionally done by precision measurements of the endpoint of the

tritium β decay spectrum. A finite $\overline{\nu_e}$ mass will distort the shape of the spectrum, most noticeably shifting the endpoint [28]. The best upper limit is currently 2.05 eV at 95% confidence level, measured by the Troisk experiment [29]. An increase in sensitivity by orders of magnitude is required to probe the expected mass ranges.

The sum of the masses of the three neutrino mass eigenstates has been constrained by various cosmological methods to be < 208 meV, to a 3σ confidence level [30].

One of the most fundamental questions unanswered about neutrinos is the origin of their mass. One possibility is that, rather than a Dirac fermion, the neutrino is a Majorana fermion. Majorana particles were first proposed by E. Majorana in 1937 [31] and are particles for which the particle and anti-particle are the same. This is possible in the case of neutrinos because they do not have charge.

Dirac neutrinos would result in four neutrinos states: left and right handed particle and anti-particle. On the other hand, Majorana neutrinos only involve two states: a left handed neutrino and a right handed neutrino, also known as an anti-neutrino [32].

If neutrinos are Majorana fermions, the PMNS matrix is altered from Eq. 1.3. U is multiplied by a diagonal matrix of $(1, e^{i\frac{\alpha_{21}}{2}}, e^{i\frac{\alpha_{31}}{2}})$ [33]. α_{21} and α_{31} are Majorana CP violating phases, just as δ_{CP} is a Dirac CP violating phase. These Majorana phases are not measurable in neutrino oscillation experiments as they are not variables in appearance or disappearance probabilities [33].

Majorana neutrinos feature in many BSM theories. They require lepton number violation [34], which can be used in combination with a seesaw mechanism to explain the expected extremely small absolute mass of the neutrino without requiring fine tuning and can be used in leptogenesis to explain the origin of the matter/anti-matter asymmetry in the Universe [32].

There is currently only one experimental method to determine if neutrinos are Majorana fermions: searches for neutrinoless double beta decay ($0\nu\beta\beta$).

1.4 Neutrinoless Double Beta Decay

1.4.1 Theory

Double β decay ($2\nu\beta\beta$), proposed by M. Goeppert-Mayer in 1935 [35], occurs in isotopes in which single β decay is not energetically allowed. Fig. 1.2 shows a cartoon of the potential β and $2\nu\beta\beta$ decays for an even A (mass number) nuclei. There are cases when β decay would involve raising the energy of the nuclei. However, $2\nu\beta\beta$ decay, (the Feynman diagram is shown in Fig. 1.3) is possible; undergoing two simultaneous β decays lowers the energy of the nuclei.
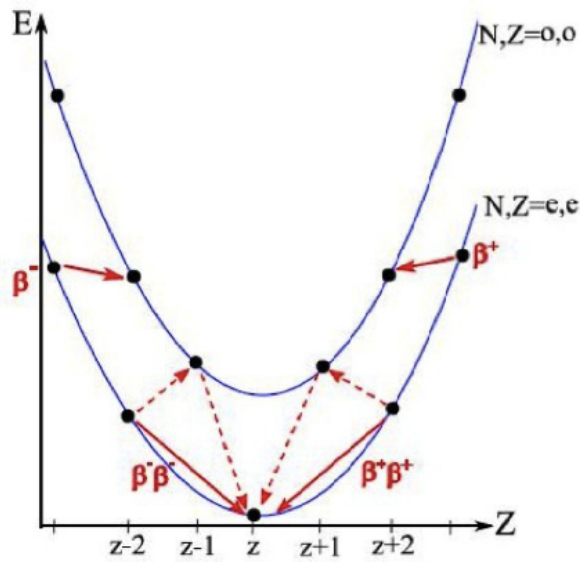
Figure 1.2: Energy parabolas (blue lines) for even A nuclei. Each black dot is an isobar (the nuclei all have the same A). Displayed in red are the possible β and $\beta\beta$ decays. Those shown with solid lines are allowed; those shown with dashed lines are forbidden. Figure from [36].
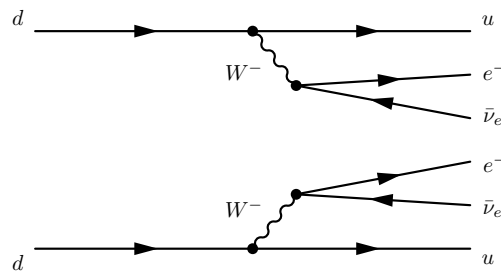
Figure 1.3: A Feynman diagram of $2\nu\beta\beta$. Figure from [55].
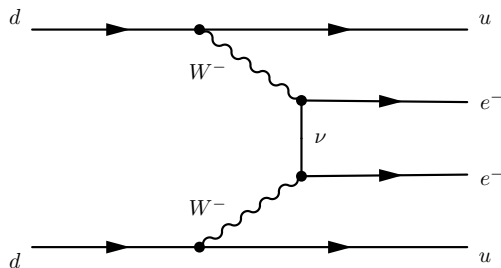
Figure 1.4: A Feynman diagram of $0\nu\beta\beta$, in this case mediated by a light Majorana neutrino. Figure from [55].

35 known, naturally occurring isotopes can theoretically undergo $2\nu\beta\beta$ [37]. However, observing it is generally only possible in nuclei where other decays (β or α) are supressed or forbiddden. It has been successfully observed in 11, with half lives between $10^{18} - 10^{21}$ years [21].

$0\nu\beta\beta$ is the hypothesis that a $\bar{\nu}_e$, emitted by one of the β decays, can be absorbed by the W boson from the other β decay as a ν_e, as depicted in Feynman diagram form in Fig. 1.4. This decay was proposed by W. Furry in 1939 [38] and has never been observed.

This process requires both that neutrinos are massive (so a helicity flip can occur) and, crucially, that they are Majorana fermions [21]. Lepton number is violated by two units, making it a BSM process.

The mass of this virtual $\nu_e/\bar{\nu}_e$ is known as the effective Majorana mass:

$$m_{\beta\beta} = \sum_{i=1}^{3} U_{ei}^2 m_i, \tag{1.5}$$

where U_{ei} is the relevant component of the PMNS matrix and m_i is the mass of the mass eigenstate ν_i [39].

1.4.2 Experiment

In $0\nu\beta\beta$, all of the energy of the decay is carried by the two electrons. The experimental signal that is searched for in experiments is consequently a peak at the endpoint of the $2\nu\beta\beta$ energy spectrum [32], as shown schematically in Fig. 1.5.

The nature of this signal makes energy resolution a critical aspect of any $0\nu\beta\beta$ experiment. Any finite energy resolution will not only spread the $0\nu\beta\beta$ peak but also smear the falling edge of the $2\nu\beta\beta$ decay energies. This can result in the $0\nu\beta\beta$ peak appearing as a shoulder to the $2\nu\beta\beta$ spectrum, or making the two indistinguishable.
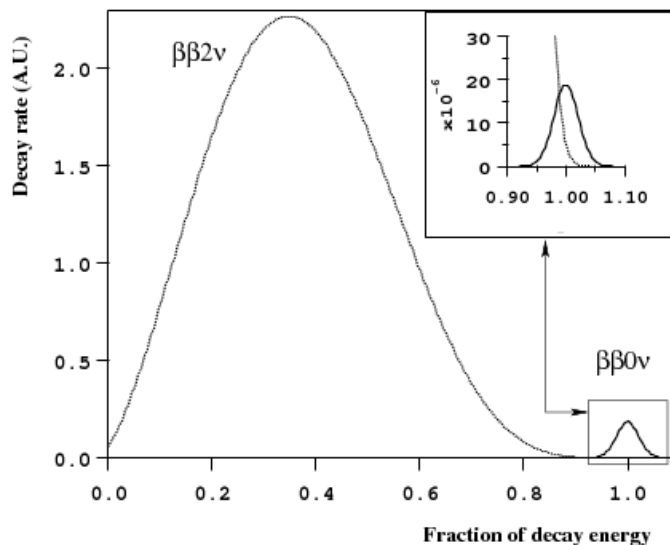
Figure 1.5: A schematic energy spectrum from a hypotheical $0\nu\beta\beta$ experiment. The signal is a peak at the end point of the $2\nu\beta\beta$ spectrum. Insert shows the same schematic, but with an increased relative number of $2\nu\beta\beta$ decays. Figure from [40].

The rate of $0\nu\beta\beta$, which is what experiments strive to measure, is related to $m_{\beta\beta}$ by:

$$T_{1/2}^{-1} = G|M|^2 \frac{|m_{\beta\beta}|^2}{m_e^2}, \tag{1.6}$$

where G is the phase space factor for the double β decay isotope, $m_{\beta\beta}$ is normalised to the electron mass and M is the specific nuclear matrix element. M cannot be calculated precisely and varies depending on the model and assumptions used. This introduces significant uncertainties into the conversion between rate and $m_{\beta\beta}$ [39]. In addition, different isotopes introduce variations in the uncertainty on M, meaning that a direct comparison between $m_{\beta\beta}$ measurements from experiments using different isotopes is challenging. However, this also motivates the use of different isotopes in experiments, in order to cancel out the uncertainty on $m_{\beta\beta}$ from M.

The dependence of the rate on the effective Majorana mass allows $0\nu\beta\beta$ experiments to probe the absolute mass scale of the neutrino. Fig. 1.6 shows the effective Majorana mass plotted against the lightest neutrino mass ($m_{lightest}$). For $m_{\text{lightest}} > 0.1$ eV, approximately, there is a degenerate region where the possible values of $m_{\beta\beta}$ in both hierarchies occupy the same region. However, below this, the $m_{\beta\beta}$ value is highly dependent on the mass hierarchy, with $m_{\beta\beta}$ values for the normal hierarchy approximately an order of magnitude, or more, lower. This means the rate of $0\nu\beta\beta$ is
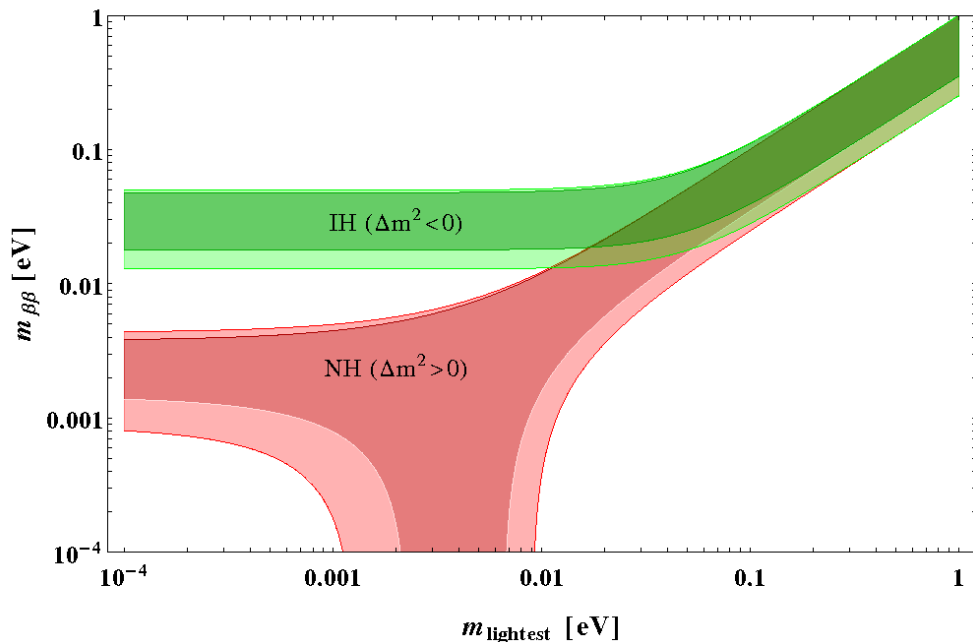
Figure 1.6: The effective Majorana mass plotted as a function of the lightest neutrino mass, for both the normal (red) and inverted (green) hierarchies. The 3σ regions and errors are indicated. Figure from [41].

correspondingly lower, making reaching the normal hierarchy regime far more experimentally challenging.

Due to this dependence, $0\nu\beta\beta$ observation or non-observation can make statements corncerning the mass hierarchy. For example, it is possible to rule out inverted hierarchy Majorana neutrinos by not observing $0\nu\beta\beta$ in experiments which have sensitivity to $m_{\beta\beta}$ as low as 0.01 eV.

As already discussed, there are benefits to searching for $0\nu\beta\beta$ in multiple isotopes. When building a $0\nu\beta\beta$ experiment, the choice of isotope is a critical decision and fundamentally affects the design of the experiment. Two of the most important factors in this decision are the Q value of the isotope (the $2\nu\beta\beta$ spectrum endpoint) and the natural abundance of the isotope.

The Q values vary between approximately 2 – 5 MeV (see Table 1.1) for the most promising experimental isotopes. In this region, specifically below 3 MeV [42], there are many natural radioactive background present. Minimising the number of events from these backgrounds (which naturally arise from the ^{238}U and ^{232}Th chains) in the region of interest for $0\nu\beta\beta$ is crucial as any $0\nu\beta\beta$ experiment is a rare event search. Therefore a high Q value is preferable.

Clearly, a large mass of the isotope is desirable. A larger mass means a greater

Isotope	Natural Abundance (%)	Q Value (MeV)
^{48}Ca	0.187	4.27
^{76}Ge	7.8	2.04
^{82}Se	9.2	3.00
^{100}Mo	9.6	3.04
^{116}Cd	7.6	2.81
^{130}Te	34.5	2.53
^{136}Xe	8.9	2.46
^{150}Nd	5.6	3.37

Table 1.1: The natural isotopic abundances and Q values of the isotopes popular for $0\nu\beta\beta$ experiments. Numbers from [43].

number of nuclei which could potentially decay via $0\nu\beta\beta$, increasing the chance of an observation. Many of the isotopes, as can also be seen in Table 1.1, have very low natural abundances, making it difficult to procure a large mass of them. It is generally possible to enrich these isotopes; however, enrichment is an expensive process.

At present, the most stringent limits on $T_{0\nu\beta\beta}^{1/2}$ come from KamLAND-Zen in ^{136}Xe $(1.07 \times 10^{26}$ yr) [44], GERDA in ^{76}Ge $(8.0 \times 10^{25}$ yr) [45] and CUORE in ^{130}Te $(1.3 \times 10^{25}$ yr) [46].

It has already been mentioned that the designs of $0\nu\beta\beta$ experiments can be extremely different based on isotope choice. For example, KamLAND-Zen is a large volume detector which uses xenon (enriched in ^{136}Xe) dissolved in liquid scintillator, GERDA is a germanium detector constructed from solid germanium enriched in ^{76}Ge and CUORE consists of cryogenic bolometers made of tellurium dioxide crystals.

SNO+, the experiment central to this work, is a $0\nu\beta\beta$ experiment which will search for a $0\nu\beta\beta$ decay of ^{130}Te dissolved in liquid scintillator. ^{130}Te was mainly chosen for is extremely high natural abundance, which means that enrichment is not necessary, combined with the relatively low cost of ^{nat}Te. SNO+, like KamLAND-Zen, is a large volume liquid scintillator detector and so the decision to use ^{130}Te was also driven by the ability to load ^{nat}Te into the SNO+ scintillator cocktail, achieving percentage level loading (by mass) while maintaining good optical characteristics [47]. This enables the detection medium to also act as the $0\nu\beta\beta$ source, meaning that a larger mass of ^{130}Te can be utilised. The SNO+ detector and scintillator cocktail is described in detail in the following chapter.

Chapter 2

SNO+ Experiment

As previously discussed, SNO+ is a general purpose, low energy neutrino experiment, with the main aim of searching for $0\nu\beta\beta$ decay in ^{130}Te [48]. This chapter will describe the detector, the detection mechanisms it uses, the three phases of the experiment and the methods of calibration utilised, as well as the software used both for simulating the detector and analysis.

2.1   The Detector

The SNO+ detector is the successor to the SNO (Sudbury Neutrino Observatory) detector and repurposes much of the equipment and experimental set-up. Therefore, it is located approximately 2 km (5890 m water equivalent) underground in the Creighton Mine, Sudbury, Canada [48] in SNOLAB, a Class 2000 clean room. This 2 km overburden acts as shielding against cosmic muons, resulting in a flux of approximately 63 muons per day through the 8.3 m radius circular area of the detector [49].

The experiment consists of a 22×34 m cavity which is sealed with urylon. This cavity is filled with $\sim$7000 tonnes of ultra-pure water (UPW), providing shielding from radioactive backgrounds originating in the rock [48]. Within this is suspended a 6 m radius spherical acrylic vessel (AV) [48] with walls 5.5 cm thick. The detection medium (discussed in Section 2.2) is located inside this AV. The AV can be accessed through an attached neck (1.5 m wide, 6.8 m long), both for the deployment of calibration sources and for filling [49]. The UPW outside the AV is referred to from here on as the external water.

The AV is supported by two high purity tensylon rope nets [48]. Due to the change in detection medium during the course of the experiment, the buoyancy of the AV varies, requiring ropes to anchor it to both the ceiling and the floor of the cavity. The two nets are called hold-down and hold-up ropes. There are 20 38 mm hold-down
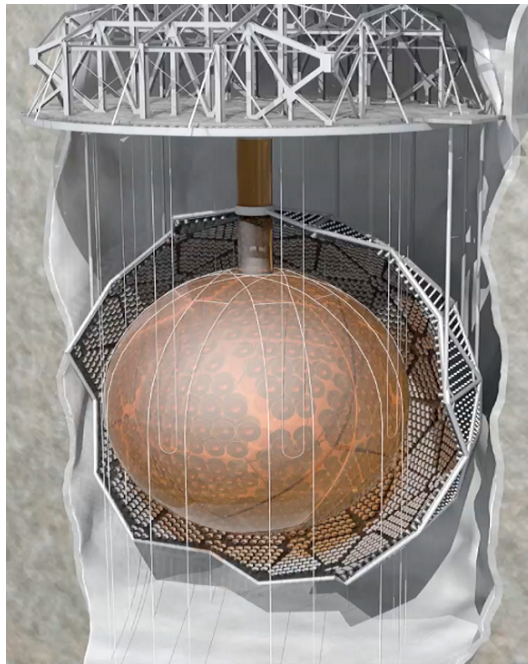
Figure 2.1: An artist's impression of the SNO+ detector. The golden sphere is the AV, surrounded by a cut away PSUP, showing the PMTs. Also shown are the rope nets. The neck of the AV connects the AV to the deck above the cavity and is used for insertion of calibration sources as well as the filling of the vessel. Figure from [50].

ropes and the 20 hold-up ropes, which were originally installed for SNO, have been replaced with 0.75 mm ones [48].

Surrounding the AV at an average radius of 8.9 m are ∼9300 Hamamatsu R1408 photomultiplier tubes (PMTs), mounted on a stainless steel support structure (the PSUP) [48]. Attached to each of these PMTs is a light concentrator: essentially a ring of curved mirrors designed to reflect light onto the cathode. This results in a PMT coverage of ∼50% [48]. The PMTs are discussed in Section 2.4 and the SNO+ DAQ and electronics are discussed in Section 2.5.

Fig. 2.1 shows an artist's impression of the SNO+ detector. For further details of the hardware, please see [48].

2.2 Detection Media

SNO+ encompasses three experimental phases, each with a different detection medium. The first phase began in May 2017, when the AV was filled with UPW. This phase was used to search for exotic physics such as invisible nucleon decay [51], resulting in world leading limits being set on proton and two dinucleon (pp and pn) partial lifetimes [51]. Additionally, the flux of ^{8}B solar neutrinos was measured [52] at en-

ergies above 5 MeV, due to low radioactive backgrounds in this regime. While this result is statistically limited, it is consistent with other experiments and demonstrates the low background nature of the detector.

For the second phase of SNO+, the AV is filled with ∼780 tonnes of organic liquid scintillator, specifically LAB (linear alkylbenzene) [48]. The filling finished in March 2021; this was delayed due to Covid-19. Combined with the LAB at a concentration of 2 g/L will be 2, 5-Diphenyloxazole (PPO), the primary fluor, and bis-MSB (1, 4-Bis(2-methylstyryl)benzene), a wavelength shifter [48]. This shifts the emitted scintillation light to better match the peak efficiency of the PMTs [53]. (The peak emission of PPO is at 359 nm, whereas the peak emission of bis-MSB is at 423 nm. The peak efficiency of the SNO+ PMT is 439 nm. See Sections 2.3 and 2.4 respectively for further details.) This scintillator cocktail will be referred to as LABPPO.

The difference in detection mechanism (discussed in Section 2.3) and background levels expected in a scintillator filled detector compared to a UPW filled detector means that there are new analyses that can be performed. Therefore, the scintillator phase will focus on the flux of low energy solar neutrinos, including CNO, pep and ^{8}B neutrinos. It will also observe the flux of reactor anti-neutrinos and perform geo-neutrino searches [54]. The scintillator phase, with its increased light output and event rate (see Section 2.3), will also fully test the detector in preparation for the final of the three phases.

The final five year phase of SNO+ is its main focus. In 2021, ∼1330 kg of ^{130}Te will be loaded into the scintillator cocktail, so that the detection medium is 0.5% ^{nat}Te by mass [47]. A search will then be conducted for $0\nu\beta\beta$ and, if $0\nu\beta\beta$ is not discovered, should provide the world's leading limit on the half life of $0\nu\beta\beta$ in ^{130}Te and the respective $m_{\beta\beta}$ (see Chapter 1) [47]. The expected energy spectrum for this final phase, with a $0\nu\beta\beta$ signal included, is shown in Fig. 2.2. This also gives an indication of the expected background levels. For further details on the $0\nu\beta\beta$ search, see [55].

Loading the ^{nat}Te is a technical challenge. High levels of metal need to be dissolved into LABPPO while maintaining transparancy and long term stability [48]. There are further requirements on the final loaded cocktail (which will henceforth be referred to as Te-loaded LABPPO), such as the ability to remove the Te once the experiment has finished and to not be flammable or a health hazard, for example. A method has been found for this loading process. Firstly, an organometallic complex of Te and 1,2-butanediol is formed. This is soluble in LAB and the product has good optical characteristics, is compatible with acrylic and can be purified by distillation [56].
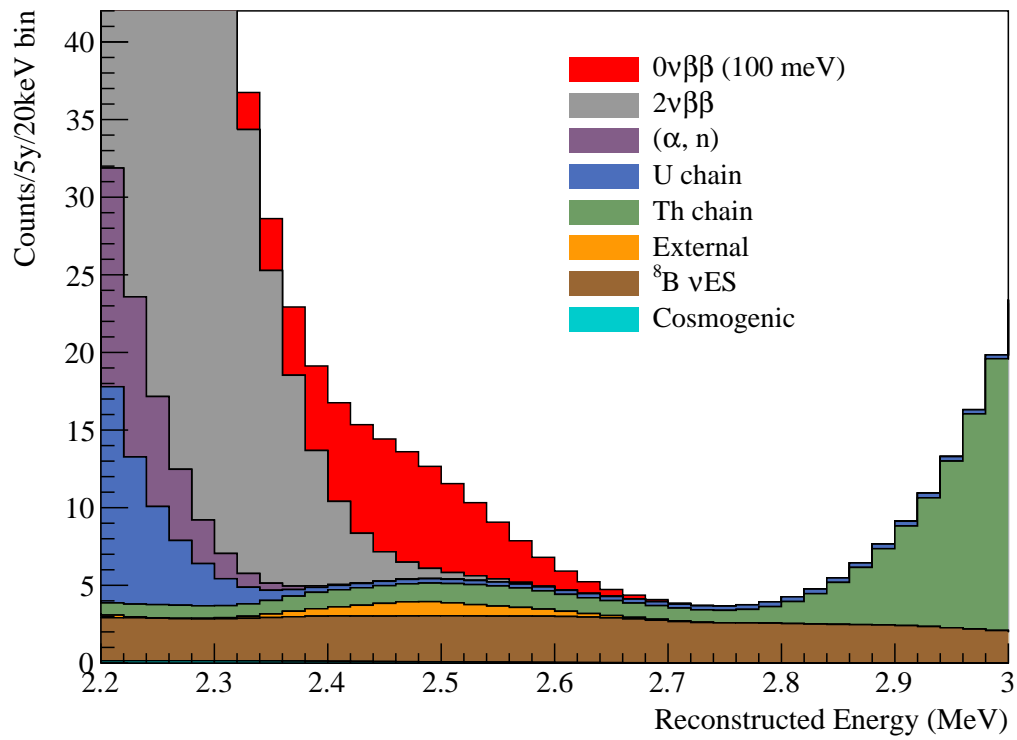
Figure 2.2: The expected energy spectrum of a $0\nu\beta\beta$ signal and backgrounds in the SNO+ detector assuming 0.5% ^{nat}Te loading by mass, a fiducial volume of 3.3 m and a run time of 5 years. Figure from [56].

2.3   Detection Mechanism

Events are detected in water and scintillator through different mechanisms. These mechanisms will only be discussed briefly in this work as neither is the source of light used in the analysis presented here.

In water, events are detected through the production of Cherenkov light. This occurs for charged particles above the Cherenkov threshold [57]. For a 1 MeV electron in the UPW filled SNO+ detector, $\sim$200 Cherenkov photons are expected to be produced. These photons are produced in a cone around the direction of travel of the charged particle. The angle of the Cherenkov radiation, θ, relative to the direction of travel of the charged particle is given by

$$\cos\theta = \frac{1}{n\beta}, \tag{2.1}$$

where n is the refractive index of the medium the particle is travelling through and β is the charged particle's relativistic correction factor [58]. Multiple scattering causes the direction of the initial particle to quickly change, meaning that this cone of photons is directionally blurred. In addition, if the charged particle is the recoil electron of a neutrino-electron elastic scattering event, the initial direction of the electron is is not necessarily the same as the neutrino. However, it is statistically correlated. Hence, in the water (or UPW) phase of SNO+, a direction can be reconstructed for an event, as well as a position and energy (the process for which is discussed in Section 2.7).

Neutral particles, such as γ's, are seen in the UPW filled detector due to processes such as Compton scattering with an electron [59].

In the scintillator phase, although Cherenkov radiation occurs, scintillation is the dominant effect [59]. In organic liquid scintillators (such as LAB), scintillation occurs due to the excitation of the π electron cloud either by ionisation and recombination or by elastic scattering. The following de-excitation to the ground state results in isotropic photon emission [60], meaning that direction reconstruction is not possible in the scintillator phases of SNO+. The wavelength (λ) of scintillation light is highly dependent on the chemical structure of the scintillator used. However, scintillation photons can be absorbed by the scintillator itself. This is partially avoided naturally by a Stokes shift [61], which causes the energy of the emitted scintillation photon to be lower than the absorbed energy. To enhance this effect, additional fluors are used to dope the scintillator. These are compounds which can absorb energy and re-emit it as photons of a lower energy. Both primary and secondary fluors are used

in the SNO+ scintillator cocktail (PPO and bis-MSB, respectively, as discussed in Section 2.2). The absorption and re-emission of photons does not just occur for scintillation photons - for the future of this work, the isotropic re-emission of photons originating from other light sources is a background that needs to be considered (see Chapter 8).

The light levels expected in the scintillator phases of SNO+ are very different than in the water phase. For a 1 MeV electron in unloaded scintillator, ~350 Cherenkov photons are expected to be produced, but ~10000 scintillation photons; the scintillator phase will have a higher light yield per MeV than the water phase. In UPW, ~8 NHits[1]/MeV are observed. In Te-loaded LABPPO, ~390 NHits/MeV are expected.

2.4 Photomultiplier Tubes

As previously mentioned, SNO+ uses ~9300 Hamamatsu R1408 PMTs, attached to the PSUP [62], via hexagonal housings. The housing also supports the concentrators, which are constructed from eighteen petals of aluminium coated in a dielectric. They are designed to increase the coverage of the PMTs by reflecting photons onto the photocathode [49]. Two diagrams of the PMTs, one including the concentrators, are shown in Fig. 2.3.

PMTs are based on the principles of the photo-electric effect and charge multiplication. The photocathode is a glass bulb coated with a caesium bialkali film [62]. When photons hit this photocathode, electrons are liberated via the photoelectric effect. These then enter the vacuum chamber behind the photocathode and, due to the photocathode being at ground potential and the anode at the base of the PMT having a high voltage (HV) of ~2000V, they are accelerated into the venetian blind style dynode stack [63]. At each collision, secondary electrons are created. By the end of the dynode stack, the charge is a measurable signal.

Fig. 2.4 shows the combined quantum and collection efficiencies of four individual R1408 PMTs, as measured in [64]. Quantum efficiency is the efficiency of creating photoelectrons (p.e.) which escape the photocathode, whereas collection efficiency is the efficiency of the p.e. to be successfully accelerated through the dynode stack.

Another important quantity associated with the PMTs is the transit time spread (TTS). The transit time is the time between a photon incident on the photocathode and the resulting current peaking. The spread or jitter in the transit times means

---

[1]Number of PMT hits, also number of triggered PMTs.
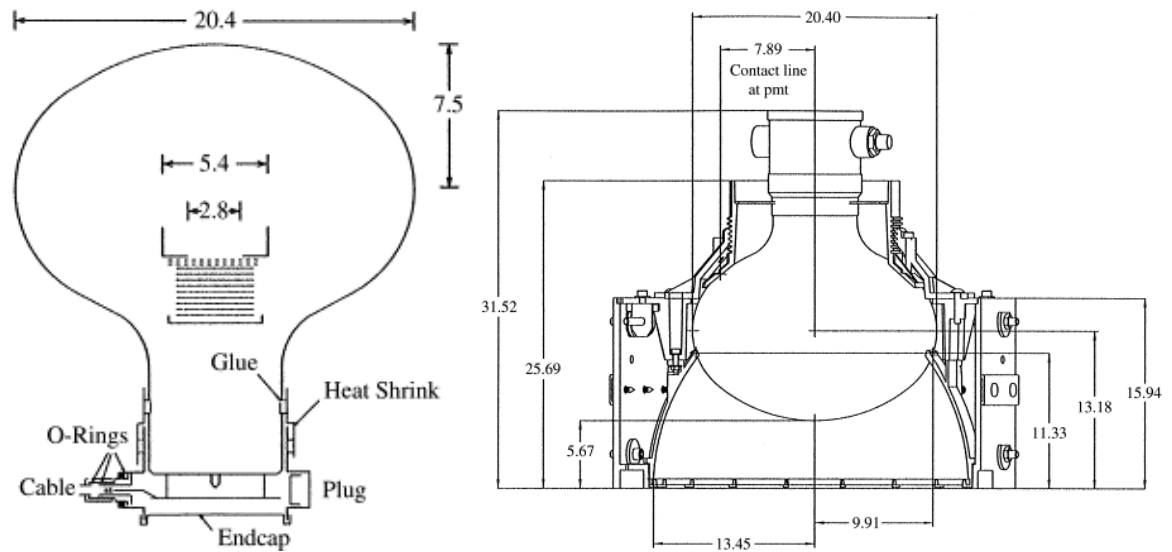
Figure 2.3: Left: A cross section of the R1408 PMT. The solid horizontal lines represent the dynode stack, the dashed horizontal line is the focusing grid and the solid 'U' is the anode [49]. Right: A cross section of a SNO+ PMT mounted in a cell with the concentrator attached [49]. All dimensions are in cm. Figures from [49].
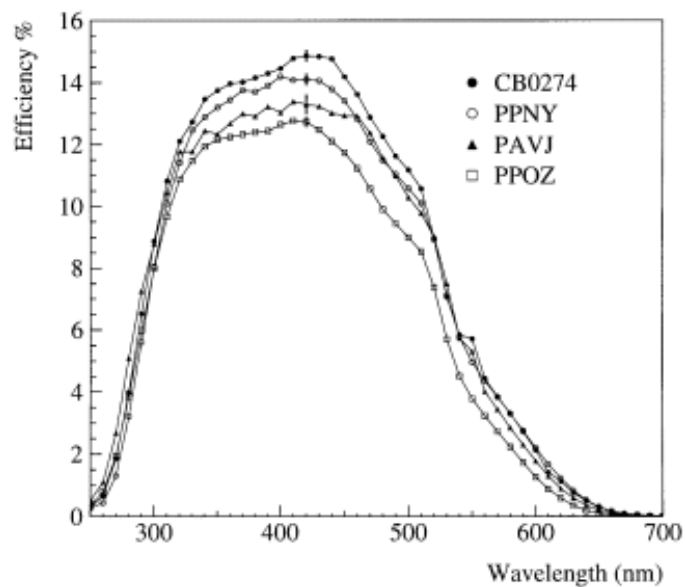
Figure 2.4: The combined quantum and collection efficiency of the SNO+ PMTs as a function of wavelength, as measured in [64]. Each line is an individual PMT. Figure from [64].

that one cannot know whether an early hit in an event was due to an early photon or a short transit time. The prompt peak of the transit time spread in the SNO+ PMTs, originally measured by P. Skensved and updated by T. Kroupova, has a full-width-half-maximum (FWHM) of ∼4 ns [65]. There is an early peak, called pre-pulsing, which is caused by photons passing through the photocathode and producing a p.e. on the dynode stack. There are also two late peaks, called late pulsing, caused by p.e. initially scattering on the dynode stack and creating secondary photons on the second time of hitting it.

As discussed above, an incident photon results in a charge on a PMT base. However, even without an incident photon there is current present - this is dark current. SNO+ integrates current (over different time durations, as discussed in Section 2.5) at the base and if the resultant charge exceeds a preset threshold for that PMT, a hit is registered. See Section 2.5 for further details of the SNO+ trigger system. This threshold is set at approximately a quarter of the average charge produced by a single electron [49].

Therefore, hypothetically, the integrated current could be used to deduce how many p.e. were present in each hit. However, fluctuations in the electron amplification process mean that the charge produced by a single primary p.e. is not a single value, but rather a spectrum. Accordingly, this is also true for any number of primary p.e.. These spectra, in the R1408 PMTs, overlap significantly at low p.e. numbers, meaning that using charge to discriminate p.e. in such a way is not a method typically utilised in SNO+ [55].

2.5   Electronics

The previous section described the mechanism by which each PMT (or channel) produces a signal pulse when it detects light. This section will discuss how this signal is processed by the detector electronics and what determines whether these signals are recorded.

The PMT signal is read out by a PMT Interface Card and passed to a daughter board (DB). The pulse is checked to see if it has crossed the discriminator threshold required to trigger that channel, as mentioned in Section 2.4. If so, the DB will create triggered pulses for that channel for each of the trigger types. The trigger signals are passed from the DB to a Front End Card (FEC), which digitises them. Next they are passed to a Crate Trigger Card (CTC), which analogue sums each type to create

per crate trigger signals. A crate is made up of the electronics for up to 512 PMTs and is physically located above the detector, on the SNO+ deck.

After the CTCs, the signals are passed to the Master Trigger Card - Analog (MTC/A). This analogue sums the signals for each trigger type, producing a set of detector wide trigger pulses. These are passed to the Master Trigger Card - Digital (MTC/D), which compares the detector wide triggers to preset thresholds. If the thresholds are exceeded within ∼400 ns of the first triggered channel, a detector wide (global) trigger is issued.

Concurrently, the DBs also measure how long the PMT signal was above its threshold. By integrating this PMT signal over different time windows, four extra per channel quantities are created. These are: timing amplitude conversion (TAC), QHS (charge at high gain over a short integration time), QHL (charge at high gain over long integration time) and QLX (charge at low gain over a long integration time). These PMT signals are sent to the FECs, along with the per channel trigger pulses, and then on to XL3[2] cards.

If a global trigger is issued, the XL3s send the time and charge information to the DAQ systems. Otherwise, they are reset. With this information from the XL3s, the event builder creates an event [59].

The global triggers are typically based on variables such as the number of hit PMTs within 100 ns - this trigger is called the NHIT100 trigger. These triggers are all synchronous. It is also possible to trigger the detector asynchronously, when the MTC/D is forced to generate a global trigger, independent of any PMT signals. This is called an external asynchronous (EXTA) trigger.

These asynchronous triggers use an auxiliary trigger board (Trigger Utility Board Mark ii - TUBii), which has been custom built to allow for an interface between optical calibration systems (the main users of asynchronous triggers, discussed further in Section 2.6.4) and the trigger system [66]. The crucial aspect of TUBii for this work is that it provides a tuneable delay on the trigger timing.

Two modes are supported for asynchronous triggers: master and slave [53]. In master mode, the trigger is issued by whichever optical calibration system is currently being deployed. This signal is sent (usually) to TUBii, which applies the requested delay, then causes the MTC/D to issue a global trigger. In slave mode, the MTC/D issues a global trigger and then a signal is sent, usually via TUBii, to the calibration system, which then fires. However, not all optical calibration systems utilise TUBii -

---

[2]Level translator boards, used for high voltage control.

some interface directly with the MTC/D - and not all of the systems can function in both modes.

Located in the same crate as the MTC/D is a waveform digitser: a CAEN v1720 [67]. This is used for digitising the global trigger signals, including the EXTA trigger. These waveforms, once digitised, are referred to as CAEN traces and are built into the triggered event [68].

Data-taking is split into runs. By default, each of these is an hour long, although the length can be manually adjusted. Each event recorded belongs to a run, with a unique run number. Crucially, the detector state during a run remains unchanged. If this is ever untrue, the run will be discarded. Runs allows the data to be split into manageable chunks and, if there is an issue in the detector, allows a minimal amount of data to be discarded. It also allows for specific runs to be marked as calibration or testing runs.

For further detail on the SNO DAQ and SNO+ updates to it, please refer to [49, 48].

2.6 Detector Calibration

As discussed, in the SNO+ detector, PMT hits are the only observable. This means that available information in an event is the number of hits, their position, time and the QHS, QHL and QLX from each PMT. The first three of these are used to reconstruct the energy and position (and direction in the water phase) of the event [48]. In order to do this, a full understanding of the detector response to different types of interaction is required. To build up this understanding four different types of calibrations are used [69], which will be discussed in the sections that follow. Calibrations are also used to validate the detector simulation and constrain systematic uncertainties on the reconstruction algorithms [70].

For these calibrations, there are two types of sources: deployed and external. Deployed sources are physically lowered into the detector through the neck using a system of ropes and can be positioned almost anywhere within the AV. However, deployed sources carry radiopurity concerns. Extensive cleaning is required before they can enter the detector and all involved systems have been manufactured to meet stringent sealing requirements [70]. Manpower underground is also required to position the sources.

A glove box (the UI - Universal Interface) is used to deploy sources. This is located on top of the neck of the AV, inside the deck clean room (DCR). On to this can be mounted the storage boxes for the ropes, umbilical (a cable used to deliver any
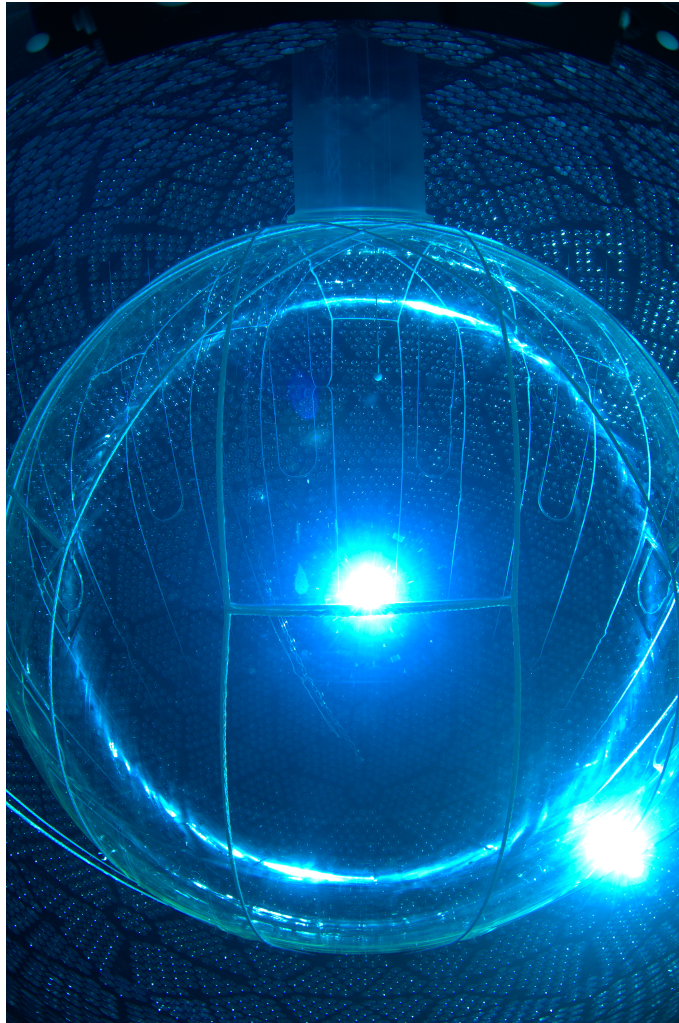
Figure 2.5: A photograph of the SNO+ detector during water phase taken by one of the cameras attached to the PSUP. External LED lamps fixed to the PSUP were used for this photograph and therefore the PMTs were on low voltage. In the top of the photograph the laserball, a deployed optical source, is visible inside the AV. Photograph courtesy of the SNO+ collaboration.

electricity, gas or other requirement for the deployed source), motors and moving mechanisms [70].

It is also possible to externally deploy sources. This means to lower them through guide tubes into the external water between the AV and the PSUP.

Additionally, six cameras have been installed on the PSUP in order to determine the position of deployed sources, as well as monitor the AV and rope positions [48]. Fig. 2.5 is an example of one of these photographs, which have been taken occansionally throughout detector operation.

External sources are optical fibres which are fixed to the PSUP (i.e. external to the AV) and cannot be moved, remaining in position for the entirety of the exper-

iment. They require less manpower to be used as they can be operated remotely, enabling much more regular and continuous monitoring of the detector response [70]. Additionally, as they are permanently located outside the AV in the external water, external sources do not pose the same level of radiation leak risk as deployed sources.

2.6.1 Electronic Calibration

Electronics calibrations (ECAs and ECALs) are the first calibrations to take place and both focus on understanding the behaviour of the readout channels of the PMTs. They are run by utilising functions built into the DAQ systems.

ECALs set the discriminator levels for each channel so that dark hits occur at a rate of 500Hz [55].

There are two types of ECA, one to calibrate the time and one to calibrate the charge response of each channel. To do this, two types of forced triggers are used: PED (the discriminator for an individual channel being forced to fire) and PulseGT (a forced detector wide trigger).

To calibrate the pedestal values of the PMT readout, PEDs are run on every channel to read out the ADC (analogue-to-digital converter) information for no hit. This will vary from PMT to PMT. This value (the pedestal) can then be subtracted when performing charge integrations. ECAs are performed regularly (usually weekly) and this means that channels which have pedestals that vary greatly over time can be flagged as inconsistent and removed from analyses [69].

ECAs are also used to relate the readout times of PMTs to the photon arrival time. Each time-to-analogue-convertor behaves slightly differently and therefore an understanding of this behaviour is needed for every channel. This is achieved by issuing PEDs and PulseGTs with varied time offsets between them [55].

2.6.2 PMT Calibration

PMT calibrations (PCAs) are designed to measure PMT related values on a PMT by PMT basis: the cable delay, time walk, angular responses, gain and relative efficiencies [71].

PCAs are conducted using either an external or a deployed source. The deployed source is the laserball, a light diffusing sphere described further in Section 2.6.4 and the external source is TELLIE, a system of 92 LEDs connected via optical fibres to injection points on the PSUP, providing wide-angle beams that cover all of the inward looking PMTs [70]. See Section 2.6.4 for further details on TELLIE.
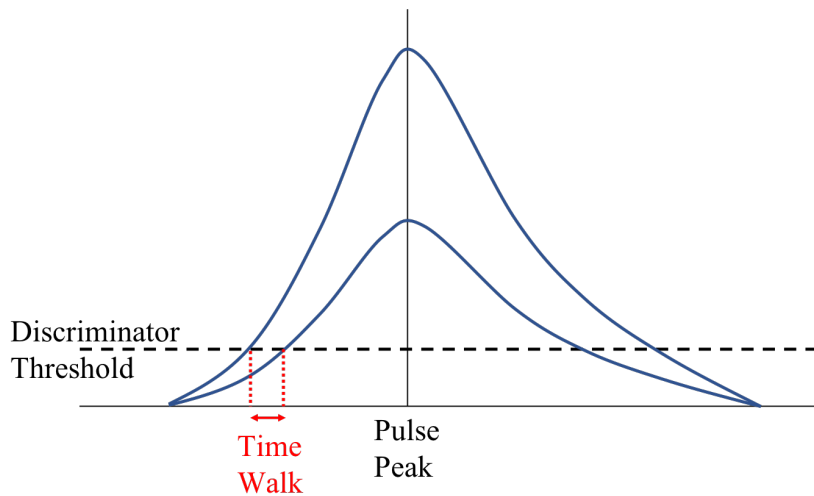
Figure 2.6: A diagram of discriminator time walk. Larger pulses, with respect to the pulse peak, will cross the discriminator threshold first. Figure adapted from [72].

Time walk describes the variation in time of the leading edge of the PMT pulse crossing the discriminator threshold. This variation arises as the time of threshold crossing is dependent on the size of the pulse. The size of the single p.e. pulse is subject to statistical fluctuations. Larger pulses will, with respect to the peak of the pulse, cross the discriminator threshold earlier [72]. This is shown in Fig. 2.6.

Cable delays and time walk are calibrated either by mapping the time versus charge space with single p.e. pulses or by using multi-p.e. pulses to measure the cable delay and single p.e. pulses to measure the time walk [73]. The requirement for time calibrations is that the uncertainty in the cable delays and time walk must be smaller than the scintillator decay time and transit time spread of the PMTs.

The gain calibration allows ADC counts above pedestal to be converted to a unit proportional to the number of p.e. [72] by characterising the single p.e. charge spectrum for each PMT. This involves extracting the threshold, peak and high-half point (HHP) for QHS, QHL and QLX from the spectrum [74]. In addition, as the gain of the PMT affects its p.e. detection efficiency, given a fixed threshold, running gain calibrations on a repeated basis allows the gain stability to be monitored [73]. This, again, allows for unstable PMTs to be removed from analyses.

Analysing the PMT hit probability with respect to the angle of the incident photons allows a model to be built of the angular response of the PMT concentrators [55].

In addition, the relative efficiencies of the PMTs can be calculated by comparing the NHits on neighbouring PMTs. These should have equal probabilities of being hit, for example, during a central laserball run, or in the centre of a TELLIE beam.

Sources	Decays (MeV)	Phases	Tagged?
^{16}N	6.1 γ	W	Yes
^{57}Co	0.122 γ	S, T	No
^{46}Sc	1.0, 1.3 γ	T	Yes
^{48}Sc	1.0, 1.1, 1.3 γ	S, T	No
AmBe	2.2 neutron, 4.4 γ	W, S, T	Coincidence tagging
Cherenkov (^{8}Li β decay)	Optical photons	W, S, T	Yes
Supernova (uses laserball)	Optical photons	S, T	Yes

Table 2.1: The isotopes and other sources used for calibrating the detector response to interactions of different types and energies. In the phases column, W represents sources deployed in the UPW phase, S sources planned to deploy in the scintillator phase and T the Te-loaded scintillator phase. Table adapted from [70].

2.6.3   Energy Calibration

The purpose of energy calibration is to understand the detector response at a range of energies and to different particle interactions. This is done by deploying radioactive sources into the detector. The decay energies of these sources are known, as is the position of the source, allowing for a measurement of NHits/MeV at various energies. This can also provide a direct comparison between simulation and data once the optical parameters of the detector have been calibrated (discussed in the next section), meaning that the performance of the position and energy reconstruction can be evaluated and constraints put upon the reconstructed quantities [48].

Table 2.1 shows the energy calibration sources that were used in either the UPW phase of SNO+, or will be used in future phases.

2.6.4   Optical Calibration

The detection mechanism of SNO+ depends on light propagation from the location of light generation (the event vertex) through the detector media to the PMTs. The aim of optical calibration is to understand all of the optical effects the photons undergo between generation and detection. This includes total attenuation, absorption, refraction, reflection and scattering, in the internal water, scintillator, Te-loaded scintillator, acrylic and external water.

In order to do this, four well understood light sources are used to inject light into the detector. These are the laserball and the three ELLIE subsystems, described below.

Laserball

The laserball is designed as a quasi-isotropic light emitter to be deployed into the SNO+ detector for both optical and PMT calibration. It was originally designed by R. Ford for the SNO experiment [71].

The laserball, shown in Fig. 2.7, is a hollow quartz spherical flask (with a diameter of 10.9 cm) filled with 0.5 kg of silicone gel. 2 g of 50 μm glass spheres filled with air are suspended in this gel. These glass spheres are uniformly distributed throughout the gel [71].

Light is emitted from a fibre which ends in the centre of the flask. The randomly distributed glass spheres then multi-refract and reflect the light, creating a quasi-isotropic far-field intensity distribution without a strong dependence on the wavelength of the light [71].

The greatest source of asymmetry in the far-field distribution is above the laserball itself, due to shadowing by the deployment hardware, as shown in Fig. 2.7. In SNO, the hardware caused a 30° shadow around the neck of the quartz flask. The redesigned SNO+ laserball (also shown in Fig. 2.7) has reduced this shadow to 7.3°. However, this redesigned laserball will only be deployed in the scintillator phases of SNO+. The redesign was necessary as the laserball now needs to be compatible with scintillator [69].

The optical fibres are contained within the laserball umbilical. These are connected to a nitrogen laser inside the DCR, which produces light at 337.1 nm with a pulse rate of 1–45 Hz [69, 71]. This laser is attached to a table with a movable mirror in its beam-line, used to direct the beam into a dye resonator. There is room for four of these resonators on the optical table. The resonators each contain a cell curvette; within each is a wavelength shifting dye. These dyes (during the UPW phase) are PBD (emits at 365 nm), BBQ (385 nm), Coumarin-450 (450 nm), BisMSB (420 nm) and Coumarin-500 (505 nm) [75]. These wavelengths have been selected to probe the Cherenkov light profile relevant for the UPW invisible nucleon decay search.

The beam-line then goes through two attenuator wheels to produce beams that have an intensity equivalent to illuminating all of the PMTs with a single p.e. [71].

Fig. 2.8 shows the entire laserball system, including the deployment equipment and laser system in the DCR.

The laserball uses the EXTA trigger in master mode, as discussed in Section 2.5. The result of this is that the global trigger always occurs with a known separation to the light injection in the detector, i.e. the time of the laser pulse is known with respect to the event trigger time [55].

Mating flange for
manipulator attachment

Umbilical
Double o-ring seal

Acrylic window

Stainless steel can

Fibre Loop

Fibre bundle - rigid fibre
connection

Restraining clamp
and light shield

Rigid fibre guide tube

Quartz flask with diffuser

Scale: 109mm
(Flask outer diameter)
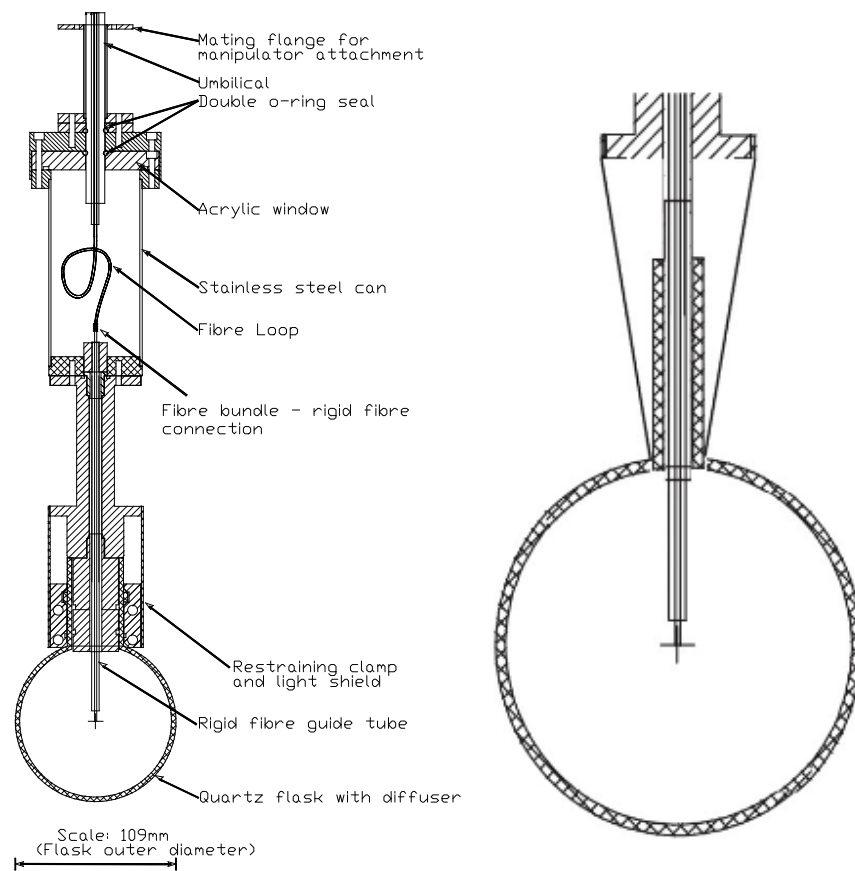
Figure 2.7: Left: A schematic diagram of the SNO laserball, a deployed optical calibration source. Figure from [71]. Right: The scintillator SNO+ laserball flask, with its differently designed light shield. Figure from [69].
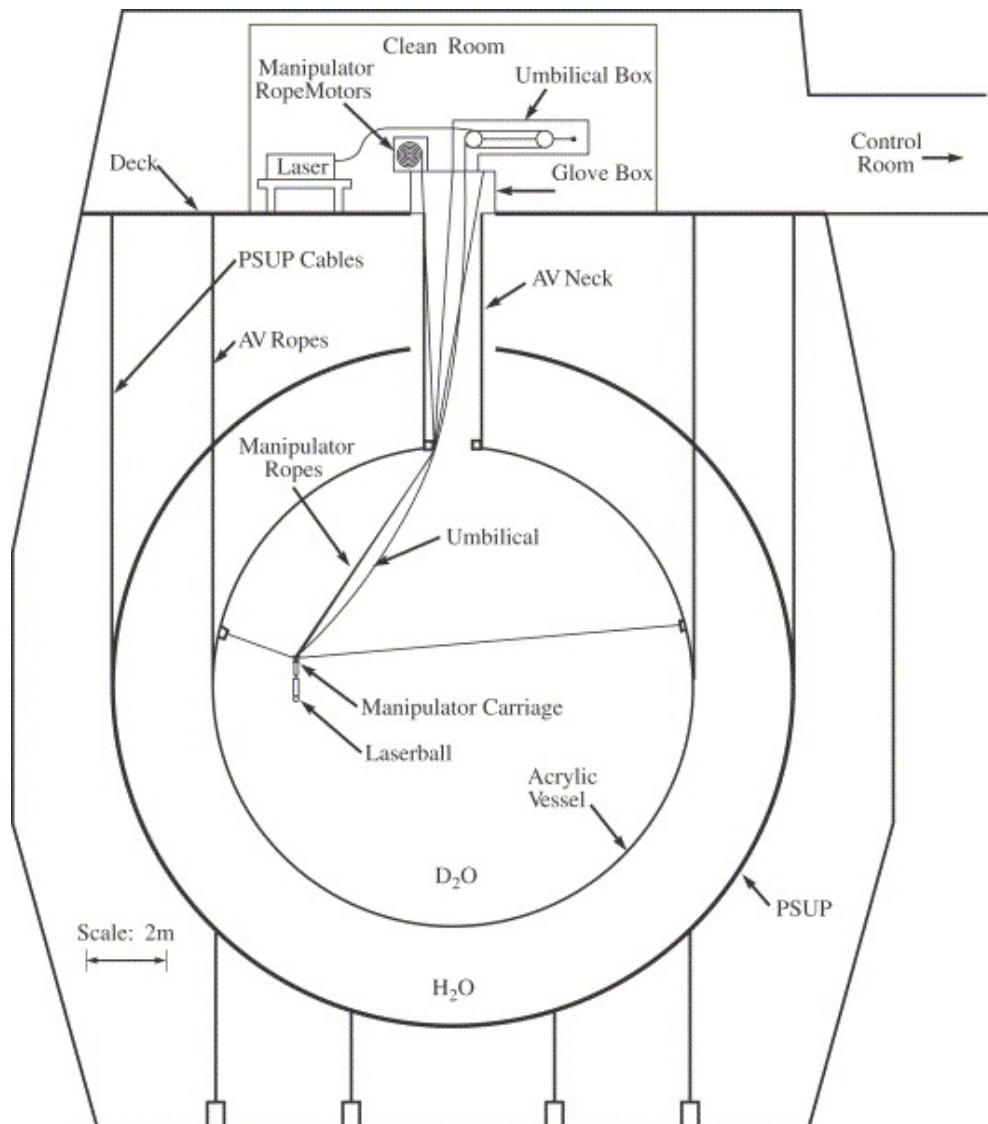
Figure 2.8: This diagram shows the laserball deployed inside the AV (of the SNO experiment) alongside all the deployment hardware. Figure from [71].

In terms of optical calibration, the laserball is used to measure, among other things, the total attenuation of the detector media (in the UPW: internal water, external water and the acrylic). In order to do this, PMT hits in a given laserball run are normalised by a run in which the laserball is deployed in the centre of the detector (called a central run), in order to remove the effects of PMT efficiencies from the analysis [76]. This ratio is compared with a prediction from the optical model of the detector (see Chapter 3). An iterative minimisation algorithm is used to fit data from a range of laserball positions (both internal and external to eliminate the degeneracy between the external water and AV effects) in order to extract the desired optical parameters for each wavelength in turn. These are then interpolated to provide values at other wavelengths. The PMT efficiencies can be extracted after this process. The parameters extracted (for the UPW phase) are: the internal and external water attenuation coefficients, the average PMT angular response and the laserball asymmetries.

There are further laserball analyses which can be performed to separately measure certain quantities: for example, deploying the laserball centrally but with varied orientations allows for the measurement of the anisotropy parameters of the laserball light distribution [76]. These results are used for initialisation of the main analysis.

The results of the laserball analysis in the UPW phase will be discussed in Chapter 3.

ELLIE

The external in-situ calibration system in SNO+ is the Embedded LED/Laser Light Injection Entity (ELLIE) [48]. It is comprised of three separate subsystems: SMELLIE (the Scattering Module), AMELLIE (the Attenuation Module), and TELLIE (the Timing module) [53]. All three of these subsystems consist of optical fibres mounted on the PSUP such that the beam of light emitted passes through the AV. These optical fibres are connected to lasers or LEDs located in the DCR [77]. Fig. 2.9 shows an ELLIE system in the SNO+ detector (with only two fibres mounted on the PSUP, for clarity).

All three subsystems use the similar mechanisms to mount the fibres on the PSUP. Fig. 2.10 shows the mechanism for SMELLIE specifically. The fibres (and collimators or terminators) plug into the holes on the holding plates. Cable ties are then used to keep the fibre straight [73].

The three subsystem also all use the EXTA trigger to ensure that the offset between the light entering the detector and the trigger time of the event is known.
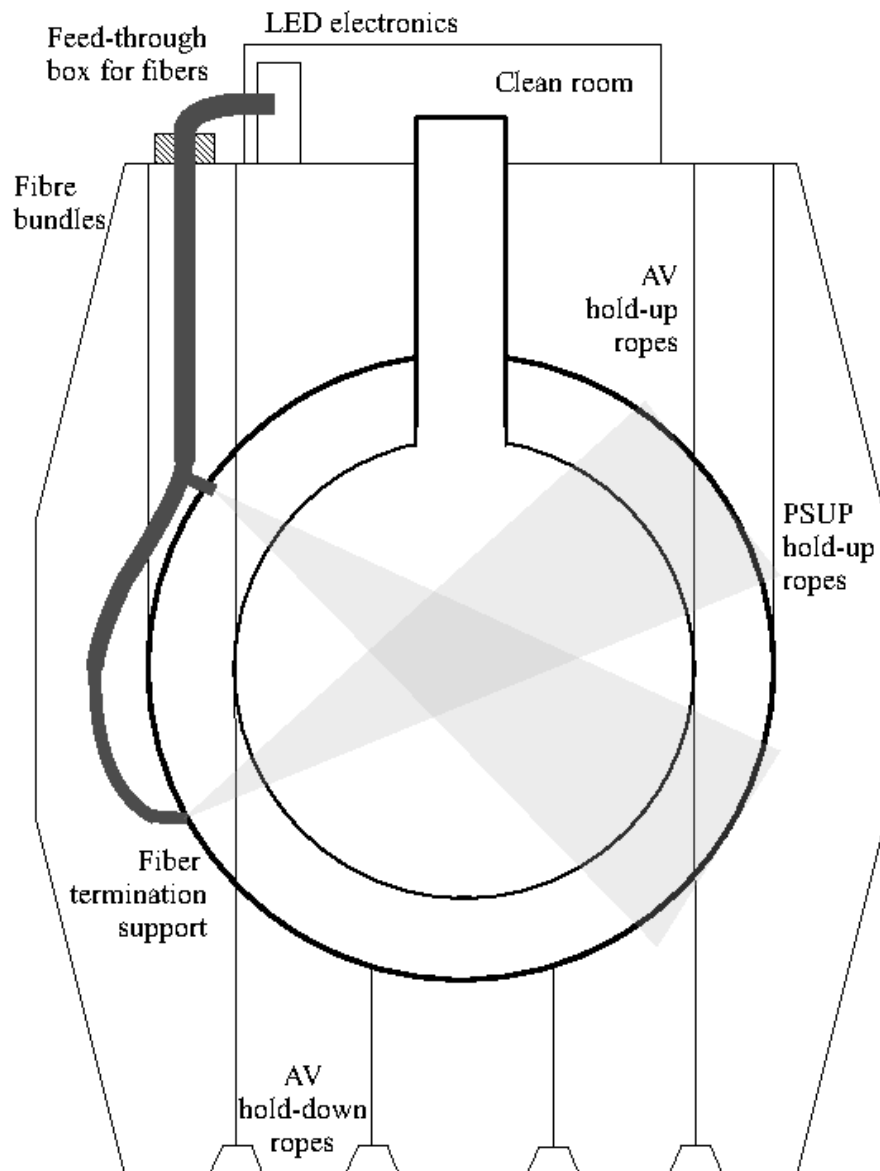
Figure 2.9: A diagram of ELLIE, an external optical calibration source which constists of LEDs/lasers on deck connected via optical fibres to injection points on the PSUP. This diagram shows TELLIE specifically. Figure from [73].

Figure 2.10: A diagram showing the mounting plates for optical fibres on the side of a PMT bucket. Figure from [73].

TELLIE is a PMT calibration system, unlike the other two ELLIE subsystems (see Section 2.6.2). TELLIE can be used to substitute for the laserball for PCAs, meaning that underground manpower is not required and without the contamination risk inherent to deployed sources. It is also used to measure the offset between the centre of the AV and the centre of the PSUP [77].

It consists of 92 injection points on the PSUP [59]. From each of these injection points a wide-angled beam is emitted, all at 0° with respect to the centre of the PSUP (in other words, pointing at the centre of the PSUP). In these beams, 80% of the light is emitted within a half-angle of 14.5° [73]. These beams cover all of the inward facing PMTs from multiple fibres.

The dry end of the fibres are connected to LED drivers of 505 nm wavelength. This wavelength was chosen in order to minimise scattering in the beams (as the probability of Rayleigh scattering decreases with increasing wavelength), while avoiding the high absorption region of Te-loaded scintillator below 450 nm and remaining as close as possible to the peak of the PMT efficiency curve at 439 nm [73].

Both AMELLIE and SMELLIE are systems for the calibration of the optical properties of the detector. SMELLIE is designed to measure the scattering characteristics of the detection media whilst AMELLIE is designed to monitor the total optical at-

tenuation length of the scintillator, the quality of the Te-loaded scintillator and warn of any degradation marked by a change in attenuation [77].

AMELLIE uses the same LED drivers as TELLIE. However, two wavelengths of LEDs can be used. Currently, the same LEDs as TELLIE are in place (505 nm), although it is planned to swap these LEDs out at a later stage of the experiment [78].

AMELLIE consists of eight narrow angle fibres at four injection points. These beams are angled at 0° and 20° with respect to the centre of the PSUP, allowing more than one path length through the detector media to be utilised [77].

SMELLIE, by measuring the scattering length, enables the attenuation length measurement to be broken down into scattering and absorption components, allowing for a more accurate simulation of the detector optics and reconstruction of events.

Unlike TELLIE or AMELLIE, SMELLIE uses lasers as its optical sources. There are five lasers: four fixed wavelength diode laser heads (375 nm, 405 nm, 445 nm and 495 nm) and one supercontinuum laser capable of producing light of wavelengths between 400 - 700 nm with a bandwidth of 10 nm [77]. This wavelength range allows for a wavelength dependent measurement of the scattering length and therefore an investigation into the nature of the scattering.

These lasers are connected via switches to fifteen collimated optical fibres at five injection points on the PSUP. At each injection point there are three beams, pointed at different angles across the detector (0, 10 and $-20°$ with respect to the PSUP centre) [77]. The SMELLIE hardware is the focus of Chapter 4 and therefore will be described in greater detail there.

Global Optical Calibration

As there are multiple sources used to measure the optical parameters of the detector, there must be a method to combine these results into a single, coherent set of parameters. Optical effects are often extremely correlated (for example, a higher probability of scattering results in a high probability of absorption, for a given absorption length, in an event window) and difficult to uniquely identify (for example, either the absorption or scattering of a photon can remove it from the temporal event window). In addition, the analyses performed for both optical sources require an initial assumption for the optical parameters. This makes combining the optical calibration results a non-trivial exercise. This combining of measurements will be discussed in practice later in this work, in Chapters 3 and 7, but will be summarised briefly here.

The laserball and SMELLIE calibration runs are initially performed independently. The laserball analysis is run with the default scattering parameters in the optical

model and assumes that any deviation from the default in the total attenuation of any detector materials is caused by the absorption component of attenuation. Using these constants, the SMELLIE analysis is done to extract a set of scattering constants. The absorption parameters are then recalculated using the measured total attenuation and scattering lengths. Following this, the laserball analysis is rerun using the SMELLIE provided scattering constants and proposes a final set of constants. The SMELLIE analysis will then be used to confirm this final set of constants. Following this, the energy calibration analyses, not described here, will be run using the updated optics.

2.7 Simulation and Analysis Software

SNO+ uses a C++ simulation and data analysis package based on RAT (Reactor Analysis Tools), originally developed by Stan Seibert [79]. It is additionally used for data processing and event reconstruction.

The Monte Carlo (MC) simulation uses `Geant4` [80] to model the full detector geometry (including PMTs, front end electronics and trigger signals) alongside a custom physics list (of relevant electromagnetic and hadronic interactions). Various generators have been written by the SNO+ collaboration to simulate a variety of physics processes, such as $0\nu\beta\beta$, solar neutrinos, radioactive background decays, etc.

The creation of optical photons from Cherenkov radiation and scintillation is performed using `GLG4sim` [81]. The optical photons are fully tracked through the detector. Each track is formed of a number of track steps. These are bounded by interactions the photon underwent, including creation, transition from one material to another, scattering and absorption. Information can be recorded for each track step including how long the step was and the two detector volumes involved in each interaction. Processes such as reflection and refraction are not explicitly recorded but are rather constructed from multiple track steps. Reflections can be identified by a photon entering then exiting a volume with a track step of length zero between the two interactions.

Finally, RATDB is a Postgres database used to store run conditions, calibration constants and reconstruction PDFs for use within RAT, both for simulation and data analysis. The default option for RAT when simulating or analysing a run is to download the relevant information from RATDB so that the the most up to date optical constants, noise levels, trigger settings, etc, for that run are used to make the simulation and analysis as accurate as possible. This is referred to as a run-level simulation.

2.7.1   Light Path Calculations

An aspect of the event reconstruction code is a class called `LightPathCalculator`, largely written by R. Stainforth [69]. This is used to estimate the probable path of a photon given a reconstructed event vertex and the location and time of a PMT hit. For two given locations, there is not one unique light path - many are possible. However, `LightPathCalculator` aims to find the shortest possible path between the two points. As well as the two points, the energy of the photon and a tolerance on the PMT position are required as arguments to the calculation. If this tolerance is set to zero, a straight line path is calculated. However, for any other value, the path takes into account refraction through the AV.

In general, the light path calculator is used in event reconstruction in order to locate the most likely event vertex. Using the group velocity of photons in the various materials and a hypothesised vertex, a quantity called the time residual (t_{res}^i) is calculated for the path. This quantity is defined as:

$$t_{res}^i = t_{hit}^i - t_{flight}^i, \tag{2.2}$$

where t_{hit}^i is the calibrated hit time of PMT i and t_{flight}^i is the flight time calculated by the `LightPathCalculator` class. A log-likelihood is then calculated using all of the time residuals from the event and a time residual PDF stored in RATDB which encapsulates detector effects (calculated using MC simulation in which the true event vertex is known). Maximising the agreement between the two sets of time residuals gives the reconstructed vertex position.

However, in the case of SMELLIE, the use of `LightPathCalculator` is somewhat simpler as the event vertex is already known; it is the fibre position. Additionally, the energy of the photon is known as the lasers fire at specific wavelengths. Therefore, `LightPathCalculator` can be used to estimate the shortest photon path between the fibre position and each hit PMT. Evidently, not all photons take the shortest path and this work is particularly interested in those photons whose path is lengthened by scattering. However, the time residual spectrum, combined with the location of each PMT hit, can be used to identify a population of hits that are highly likely to have been caused by scattered photons. This will be discussed in much greater detail in Chapters 5 and 7.

Of particular interest when using `LightPathCalculator` with SMELLIE events is functionality provided by the `SetELLIEReflect` boolean [82]. This is designed for use with light sources that are external to the AV (typically ELLIE, but also used

for external deployment of the laserball, for example). When PMT hits are in the geometrical region in which a straight line from the fibre position to the hit PMT does not intersect the AV, reflections from the AV surface are considered as possible light paths. The use of this boolean will be discussed further in Chapter 5.

Chapter 3

Optics

3.1 Theory of Optical Scattering

This work aims to measure the scattering characteristics of the SNO+ detection medium. Therefore, the first section of this chapter will discuss the theory of optical scattering and the characteristics of particular types of scattering.

Optical scattering refers to the scattering of optical photons off of particles. This process, together with absorption, comprises the attenuation of a beam of photons, otherwise known as extinction [83]. While the attenuation of the detector medium is measured by other optical calibration systems, discussed in Chapter 2, one of the fundamental purposes of SMELLIE is to enable the decoupling of the absorption and scattering components of the measured attenuation.

Rather than discussing all scattering, this work is only concerned with elastic scattering. Inelastic scattering, when the energy and frequency of the scattered photon is different to the incident one, is known as Raman scattering and is not considered here due to its extremely low occurrence rate in nature [84].

Optical scattering is an electrodynamic process. The general case of the scattering of an electromagnetic plane wave by a homogenous sphere is described by Mie theory [85]. However, in the case of the scatterers being much smaller than the wavelength of the photons being scattered, Mie theory is approximated by Rayleigh scattering [86]. In the opposite regime, when the size of the scatterer is significantly larger than the wavelength of the incident photon, Mie theory converges to geometric optics [83]. Mie scattering typically refers to the regime in between these two extremes.

This work is primarily focused on Rayleigh scattering as this is the type of scattering expected to be dominant in the detection media of SNO+ [87]. Understanding

the characteristics of Rayleigh scattering is vital, therefore, to evaluate whether observations match expectations, and, if not, why this might be. Therefore, as a brief overview of the general case, Mie theory will be discussed next, followed by further detail on Rayleigh scattering.

3.1.1 Mie Theory

Mie theory, created by Gustav Mie in 1908 [85], describes the elastic scattering of a photon off of a spherical particle. Scattering occurs when the orbital electrons of the scatterer are displaced by the electromagnetic field of the photon. This causes the electrons to move in an oscillatory motion with an angular frequency (ω) the same as the frequency of the electromagnetic field of the photon. This oscillation creates an induced dipole moment on the scatterer. The radiation emitted by this dipole is the scattered light [88].

In the general case, the differential cross section for Mie scattering of unpolarised light is an infinite series of Legendre polynomials [88]. Solving these is typically done using computational algorithms specifically designed for the task [89].

3.1.2 Rayleigh Scattering

As previously mentioned, the general case of Mie scattering can be approximated by Rayleigh scattering. For this to be applicable, the two Rayleigh criteria must be met. These arise from the requirement that the wavelength of the photon is much larger than the size of the scatterer - this means, for one, that the electromagnetic field of the photon is uniform as far as the scatterer is concerned. It also means that the time period in which the induced dipole moment is established is shorter than the electromagnetic field oscillation period of the photon [83].

The differential Rayleigh scattering cross section that arises from considering Rayleigh scattering as a limiting case of Mie theory is

$$\sigma'_{RS}(\theta, \phi) = \frac{8\pi^4 a^6}{\lambda^4}(\frac{n_{par}^2 - 1}{n_{par}^2 + 2})^2(1 + \cos^2\theta), \tag{3.1}$$

where λ is the wavelength of the photon in the bulk medium, a is the radius of the scatterer and n_{par} is its refractive index (with the assumption that the refractive index is real). θ is the polar angle and ϕ the azimuthal angle of the scattered photon with respect to the incident photon [88].

Crucially to this work, the differential cross section has a λ^{-4} dependence. This is a notable feature of Rayleigh scattering and is used later in this work to test whether

the observed scattering fits the behaviour of Rayleigh scattering or whether it is caused by another phenomenon. It is also important to note the $\cos^2 \theta$ dependence of Rayleigh scattering as this provides a mechanism to distinguish scattering from re-emission, a process which, in the SNO+ detector, shares very many characteristics with scattering but is isotropic. This will be further discussed in Chapter 8, as it is not relevant for the water phase, but will be for future phases.

The Raleigh scattering length (l_{RS}), a core concept in this work, is defined as the mean distance a photon travels in the medium before it Rayleigh scatters. This is the quantity measured by SMELLIE as a function of wavelength in Chapter 7.

So far, this discussion has been relevant to ideal materials rather than bulk ones. Treating ideal materials means that the scatterers can be treated in isolation. Therefore, inevitably, Eq. 3.1 does not strictly hold for any real materials due to forces between the particles in the material. One of the more major differences arises from the electric polarisability; an ideal material is optically isotropic. This is not true in bulk materials - they are anisotropic. This anisotropy is captured in a quantity known as the depolarisation ratio (δ) and introduces a factor of $\frac{1-\delta}{1+\delta}$ on the $\cos^2 \theta$ in the differential Rayleigh scattering cross section [53].

The implementation of δ in RAT is not discussed here, but further reading is available in [53]. A discussion on the implementation of scattering more generally follows.

3.1.3   Implementation in RAT

The simulation of the paths of optical photons has already been discussed, in Section 2.7. In addition to the functionality provided by GEANT4, a custom process is used specifically to determine if a photon undergoes attenuation called `OpAttenuation`. This process takes into account the total attenuation length of the photon, which is wavelength dependent. Before a track is created for the photon, a random number is thrown from a exponential distribution of the total attenuation length in order to calculate the length of the next track step. (The track steps are straight lines from the position of the photon along its direction vector.) If the straight line distance to the next volume boundary (such as that between the external water and the AV) is shorter than the thrown track length, the track step will end at the boundary. In this case, refraction or reflection will occur, depending on the refractive indices of the materials that make up the two volumes. A new track step is then created to propagate the photon in its new direction [59].

However, if the straight line distance to the volume boundary is longer than the track length thrown, the track step will end after the thrown distance. As the thrown distance was drawn from the exponential distribution of the total attenuation length, it is already determined that attenuation occurs. However, it has to be decided whether this takes the form of absorption or scattering. To do this, the scattering fraction is calculated as $\frac{l_{total}}{l_{RS}}$, where l_{total} is the total attenuation length. A random number is thrown from a uniform distribution between 0 and 1 and, if it is lower than the scattering fraction, the photon is determined to have been scattered. Otherwise, it will be marked as absorbed, which stops the tracking process for that photon [59].

If the photon is scattered, a new direction for the next track step of the photon track will be thrown from relevant angular scattering distribution.

Scattering in the detector simulation is only modelled as Rayleigh scattering. Mie scattering is not simulated. Therefore, the angular distribution used to throw the new photon direction for a scattering event is that of Rayleigh scattering: $1 + \cos^2\theta$. This is assumed to be a valid approach as, in all measurements of pure water, Rayleigh scattering has been shown to be the dominant scattering process [90]. This is also true for measurements of LABPPO and the Te-loaded LABPPO measurements that have been conducted by the SNO+ collaboration [87].

3.2 Optical Properties of the Detector

Relating the measured optical properties of the detector to those in the simulation is a key theme of this work. Therefore, a further discussion of the optical properties in the detector simulation is required.

The detector simulation is formed of GEANT4 volumes. Each of these are constructed from materials, which are constructed out of elements [80]. These materials are assigned various optical properties. Volumes, as well as having the bulk properties of the materials that form them, can be assigned surface properties.

The optical properties assigned to these volumes need to be ones which can be measured and accordingly altered in the simulation [59]. Exactly what properties are assigned depends on the material in question. For example, light yield, the emission spectrum and timing must be included for the scintillator cocktails. In the UPW phase of SNO+, however, these properties are clearly not required and therefore not assigned. As this work is focused on the calibration of the UPW phase, only the properties relevant to this phase will be discussed here.

These properties are:

- Scattering length,

- Absorption length,

- Refractive index,

- Group velocity.

The need for accurate measurements of the first two of these properties has been discussed in Chapter 2, as has the method of combining the laserball and SMELLIE measurements in an iterative process to measure them.

Knowledge of the refractive indices of the various detector media allow for accurate simulations of reflection, refraction and are required for the calculation of group velocities. The refractive indices used for the media in the water phase of SNO+ were taken from a combination of theory and ex-situ measurements performed both for SNO and SNO+.

The group velocity constants allow for conversion between photon travel distance and travel time. They are able to be measured using laserball data. This was done using water phase laserball data from December 2017 and the values found were entirely consistent with those already in RATDB. Therefore, they will not be discussed during this work. See [76] for further details.

All of the four properties are measured as a function of wavelength and are recorded in RATDB tables for each material over the wavelength range of interest (200 - 800 nm, though parts of this range may be extrapolated as a constant). For the absorption and scattering lengths, a scaling factor is introduced in the simulation, so the magnitude of the scattering or absorption can be easily adjusted without needing to re-calculate the values for each wavelength. Both of these scaling factors are divisive, meaning that an increase in the scaling factor leads to a decrease in the absorption/scattering length and therefore an increase in scattering/absorption. The scaling factor to be applied to the scattering length is the main quantity which this work sets out to measure.

In this work, there are three different optical models of the detector that are considered. These consist of different combinations of absorption and scattering lengths used for the internal water, external water and acrylic, arising from measurements made at different stages of the water phase calibration campaign.
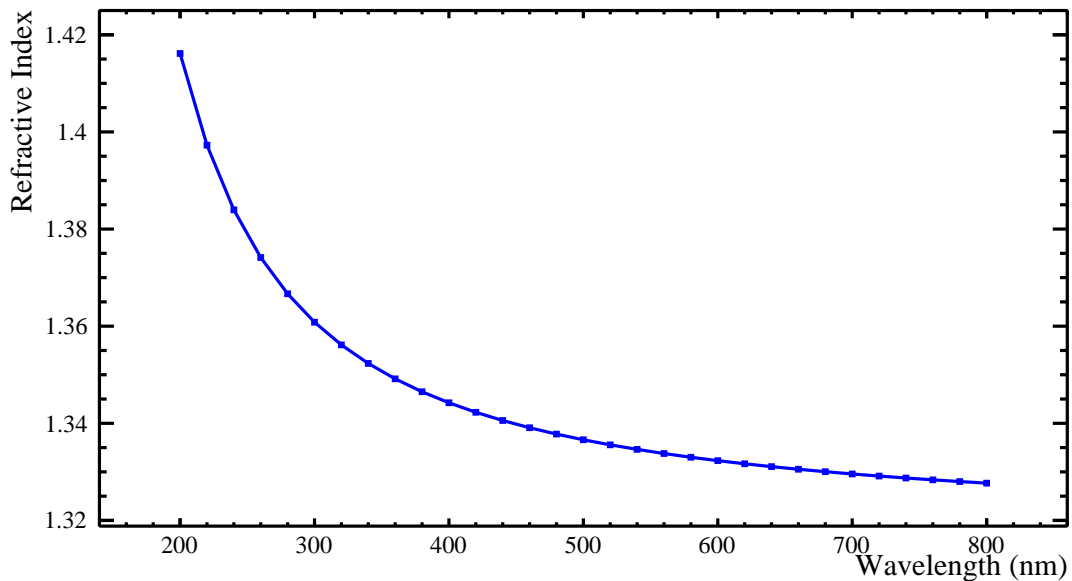
Figure 3.1: Refractive index of internal water as a function of wavelength as in the optics tables in RATDB at the start of SNO+ water phase.

3.2.1 RAT Optics

The first set of optics to be discussed are the optical constants that existed in RAT prior to the calibration campaign in water phase. These will henceforth be referred to as the RAT optics. These were largely implemented from an extrapolation of the values used for SNO [76].

Figs. 3.1 – 3.3 show the refractive index, absorption length and scattering length of the internal water as in RATDB at the start of this work.

The values shown for l_{RS} in Fig 3.3, as discussed in [53], are not entirely consistent with theoretical calculations or values in literature. This is not unexpected as, during the derivations of these values from SNO constants, assumptions had to be made regarding the split of the attenuation length into its scattering and absorption components. However, at the wavelengths of interest in SNO+, l_{RS} itself is much longer than the diameter of the detector. Therefore, as one of the aims of this work is to improve upon this initial set of optical constants, these differences do not invalidate using the shown constants as the initial set for the SNO+ simulations.

Figs. 3.4 – 3.6 show the same respective values for the external water and Figs. 3.7 – 3.9 for the acrylic of the AV.
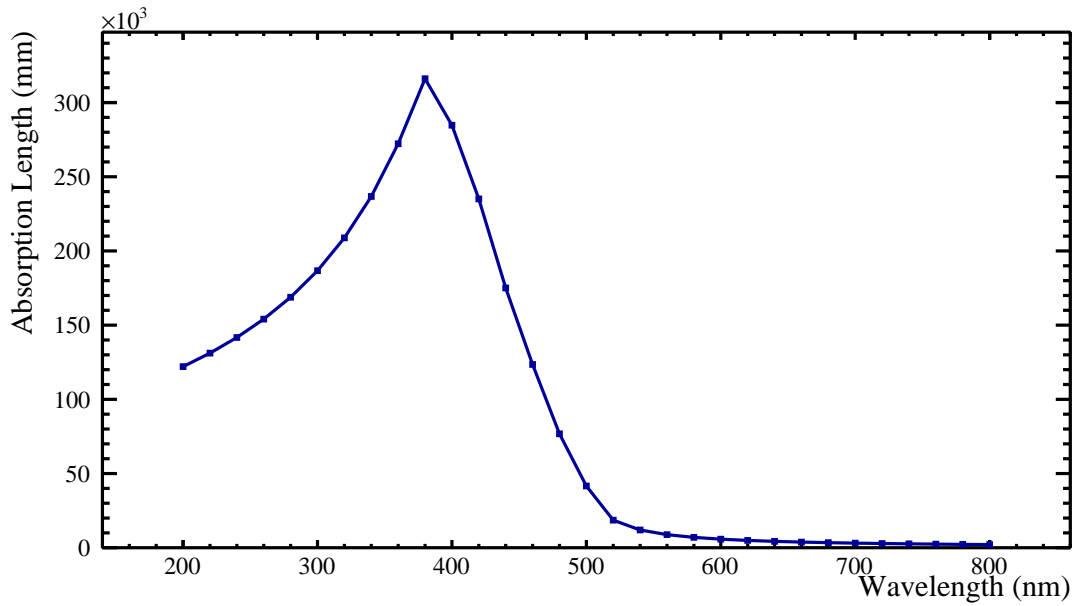
Figure 3.2: Absorption spectrum of internal water as in the optics tables in RATDB at the start of SNO+ water phase.
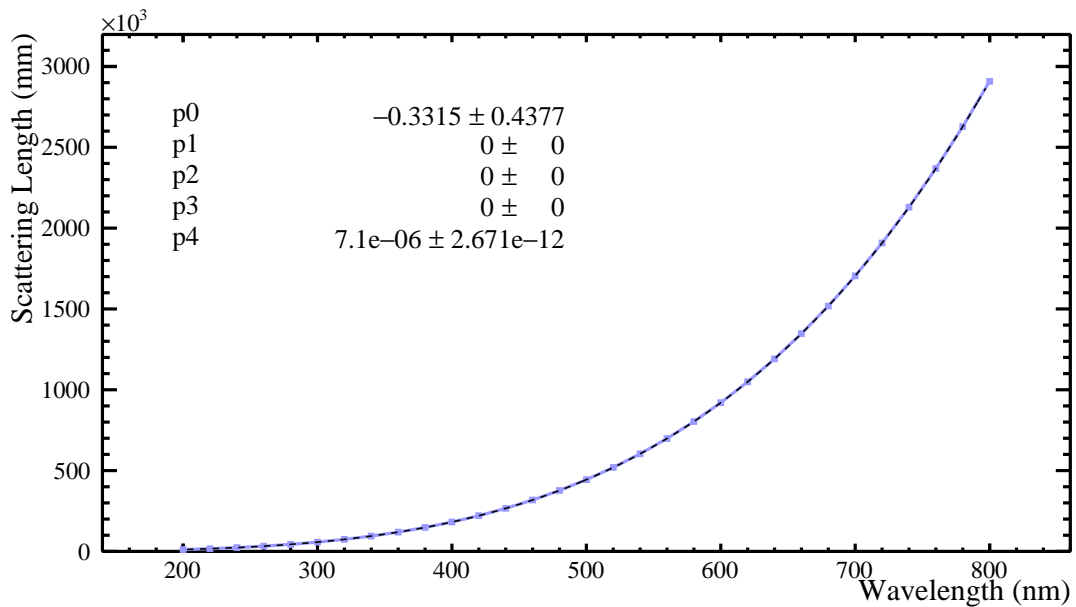
Figure 3.3: Scattering length of internal water as a function of wavelength as in the optics tables in RATDB at the start of SNO+ water phase. The spectrum is fitted with a function of the form $y = p0 + p1\lambda + p2\lambda^2 + p3\lambda^3 + p4\lambda^4$ (the dashed line) where p1, p2 and p3 are fixed to zero to match the expected Rayleigh scattering wavelength dependence.
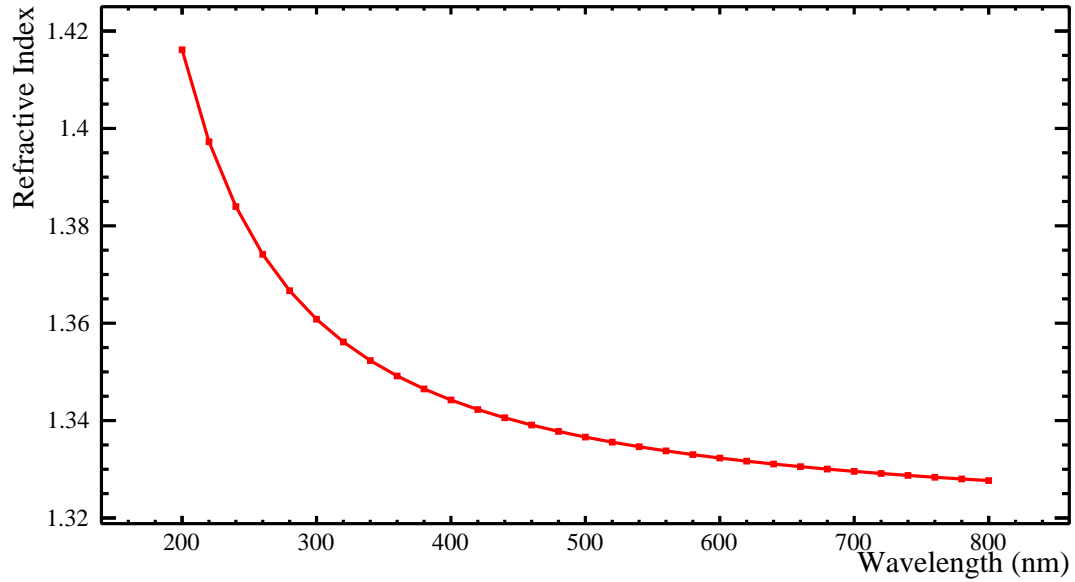
Figure 3.4: Refractive index of external water as a function of wavelength as in the optics tables in RATDB at the start of SNO+ water phase.
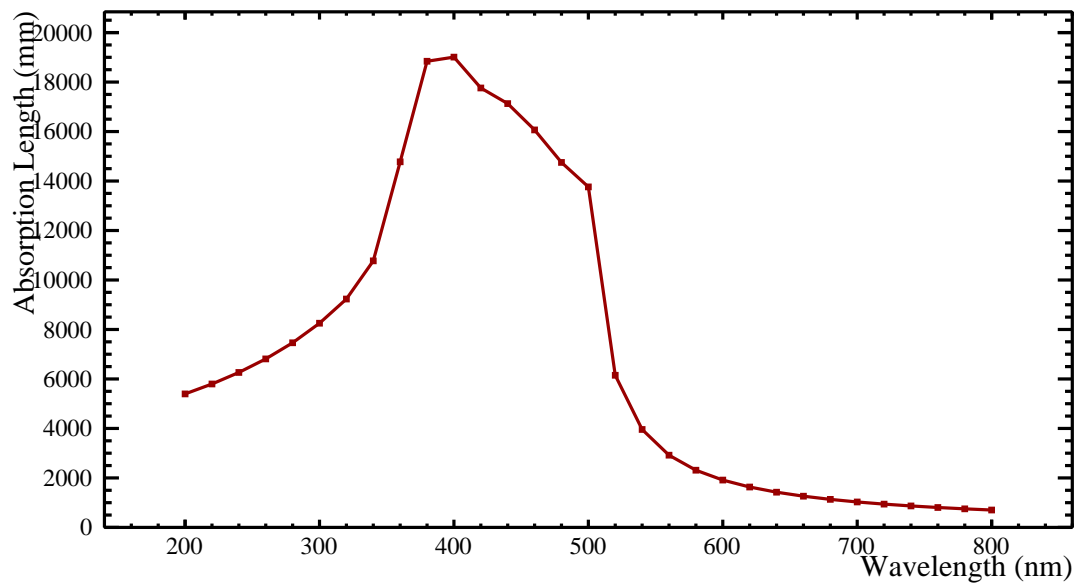
Figure 3.5: Absorption spectrum of external water as in the optics tables in RATDB at the start of SNO+ water phase.
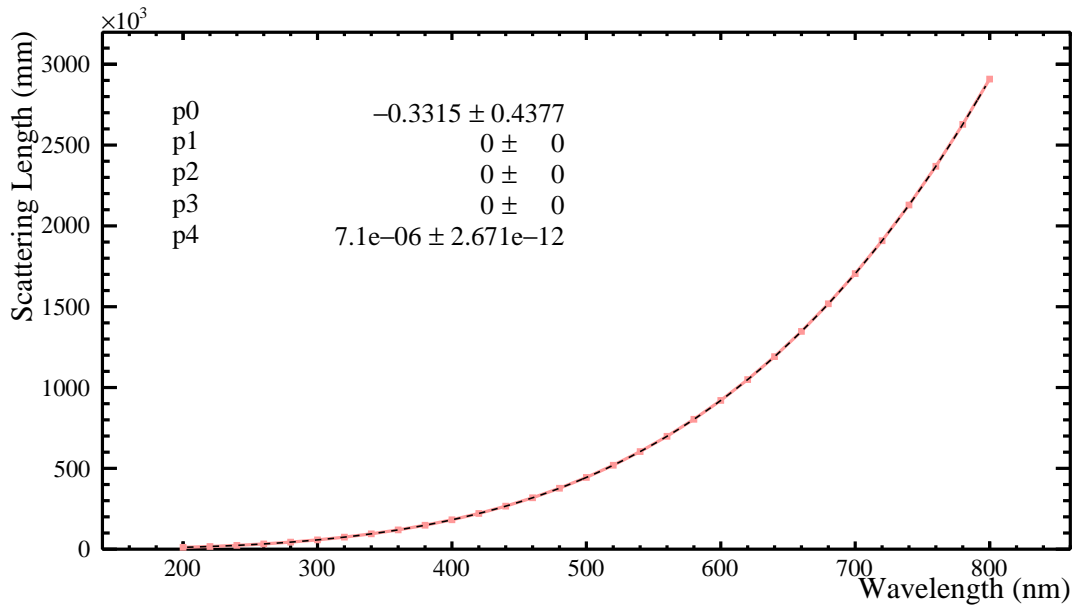
Figure 3.6: Scattering length of external water as a function of wavelength as in the optics table in RATDB at the start of SNO+ water phase. The spectrum is fitted with a function of the form $y = p0 + p1\lambda + p2\lambda^2 + p3\lambda^3 + p4\lambda^4$ (the dashed line) where p1, p2 and p3 are fixed to zero to match the expected Rayleigh scattering wavelength dependence.
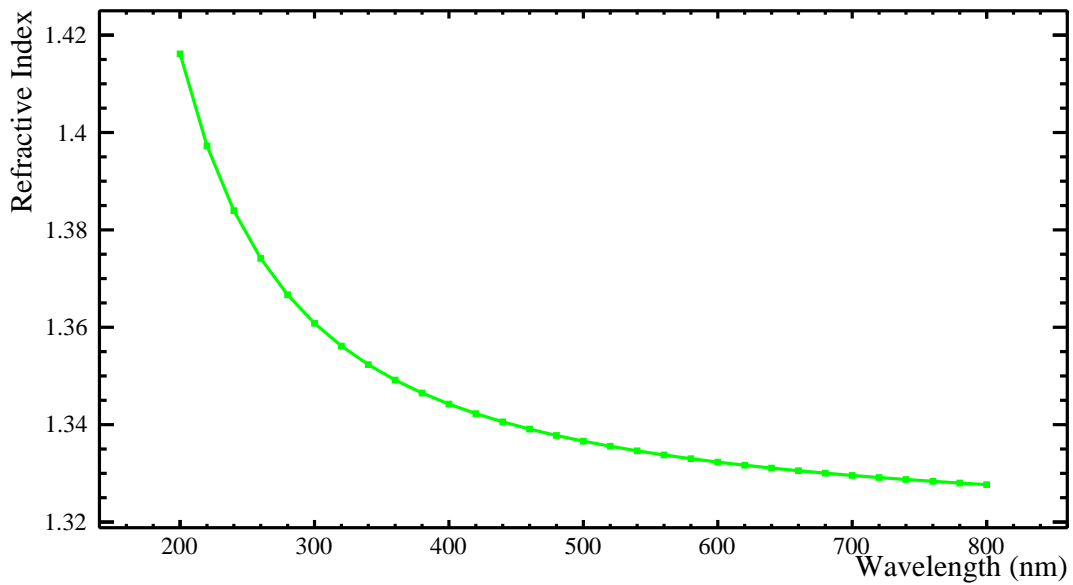
Figure 3.7: Refractive index of acrylic as a function of wavelength as in the optics tables in RATDB at the start of SNO+ water phase.
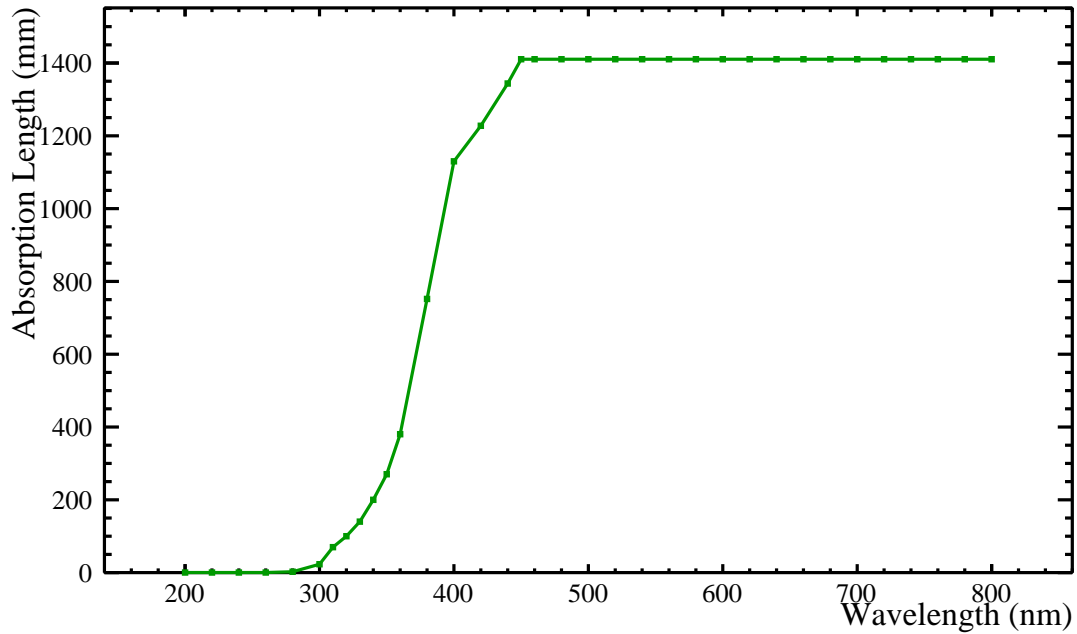
Figure 3.8: Absorption spectrum of acrylic as in the optics tables in RATDB at the start of SNO+ water phase.
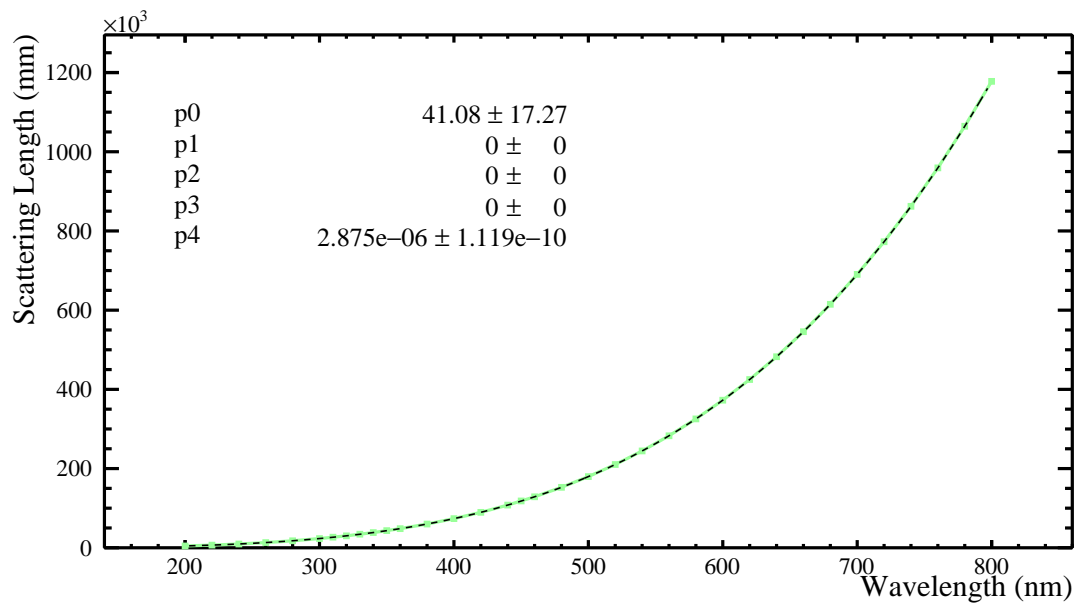
Figure 3.9: Scattering length of acrylic as a function of wavelength as in the optics tables in RATDB at the start of SNO+ water phase. The spectrum is fitted with a function of the form $y = p0 + p1\lambda + p2\lambda^2 + p3\lambda^3 + p4\lambda^4$ (the dashed line) where p1, p2 and p3 are fixed to zero to match the expected Rayleigh scattering wavelength dependence.

3.2.2 Proposed Optics

During the course of the first SMELLIE analysis performed in this work, a new set of optics constants were proposed (henceforth called the proposed optics). These arose from the analysis of laserball data taken during the water phase (mainly internal and external water scans taken in July 2018). Using scans with the laserball deployed in the external water is of particular interest as it allows the correlation present in the internal scans between the attenuation lengths of the external water and the acrylic to be broken [76]. This is because, in external laserball runs, the light does not have to traverse the acrylic to reach the PMTs closest to itself.

During the analysis of these runs, a decision was made to assume the same attenuation coefficients for both internal and external water. While the laserball analysis did not originally make this assumption, it was found that the two sets of coefficients were compatible with each other (although with larger uncertainties on the external water coefficients) [91].

This, in turn, causes the attenuation length of the internal water to be slightly lowered from its initial value in RATDB. Until the results from the SMELLIE analysis were available (which was not the case for this set of optical constants), the scattering length was kept constant in all water volumes, so the change in attenuation was attributed to the absorption length [92]. The initial RATDB and newly proposed absorption lengths for internal and external water are shown in Figs. 3.10 and 3.11.

Attention then turned to making an in-situ measurement of the attenuation of the AV acrylic for the first time from the same laserball data. To account for the change in the attenuation length in the waters (particularly the more dramatic change in external water), the attenuation length of acrylic was decreased [76].

As with all materials in RAT, the attenuation of acrylic in the simulation is split into absorption and Rayleigh scattering spectra. As the acrylic is a very thin medium, separating the additional measured attenuation into these two components is not trivial. This was attempted by running central laserball simulations at 420 nm with varied amounts of the extra attenuation attributed to scattering and the rest to absorption. The scattering lengths in these simulations were multiplied by factors between 0.0045 and 0.01 (factors larger than this resulted in the attenuation being dominated by absorption) as well as a factor of one [76].

As a result of these simulations, a multiplicative factor of 0.01 (equivalent to a scaling factor of 100) was initially chosen to be applied to the scattering length of acrylic, with the remainder of the extra attenuation attributed to absorption. This was due to suspicions that activity during the transition from SNO to SNO+ could
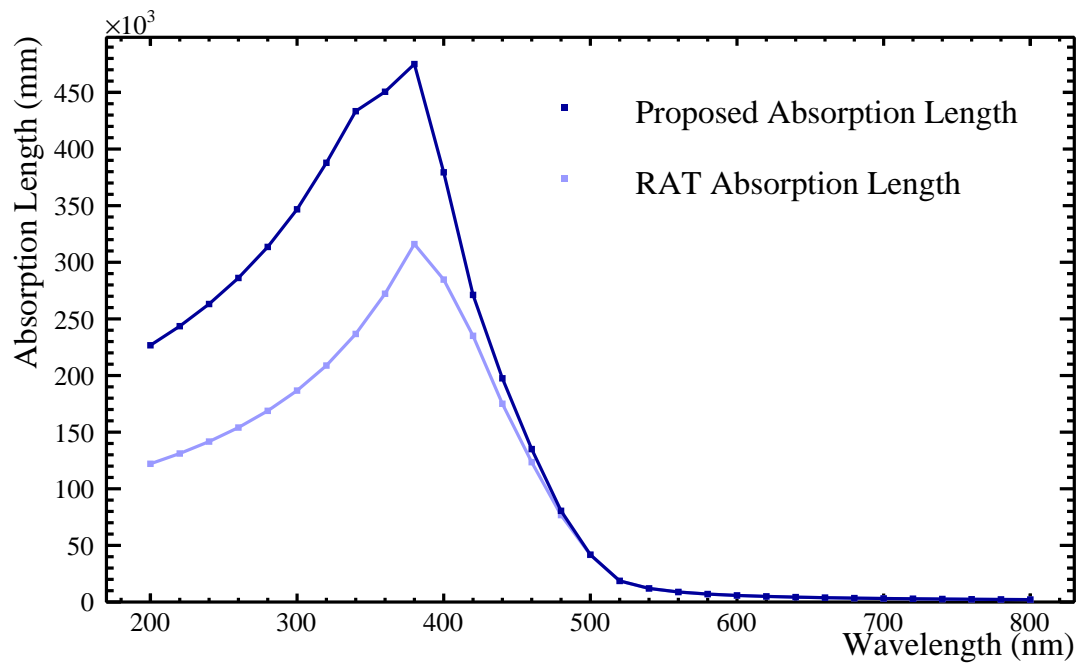
Figure 3.10: Two absorption spectra of internal water: in the light blue is the initial spectrum from RATDB, in the dark blue is the spectrum proposed after the laserball analysis.
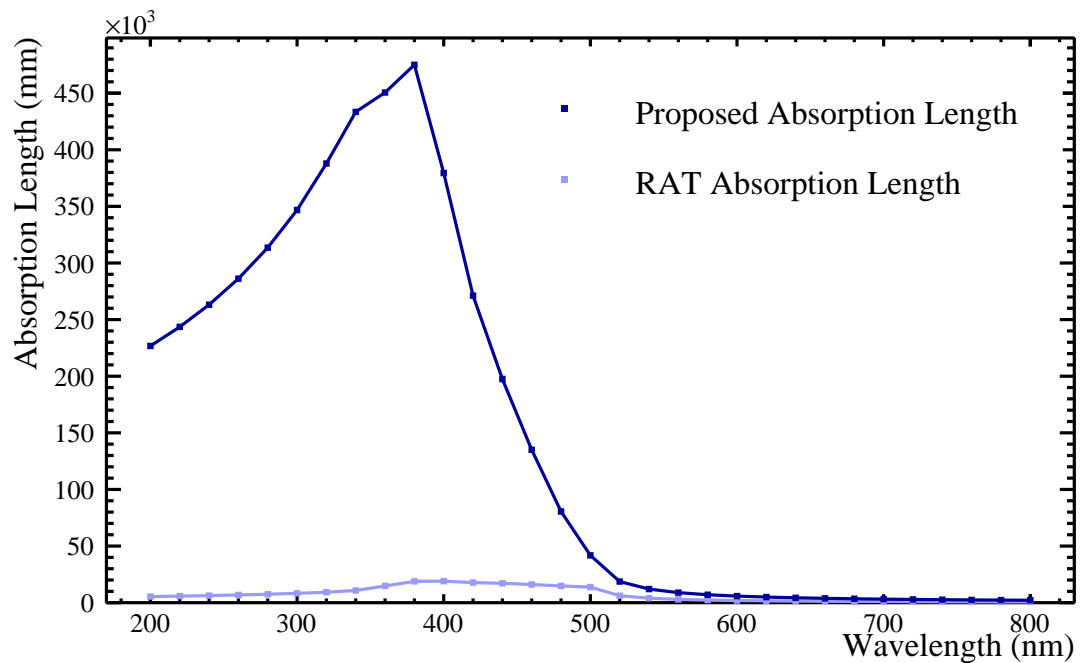
Figure 3.11: Two absorption spectra of external water: in the light blue is the initial spectrum from RATDB, in the dark blue is the spectrum proposed after the laserball analysis.
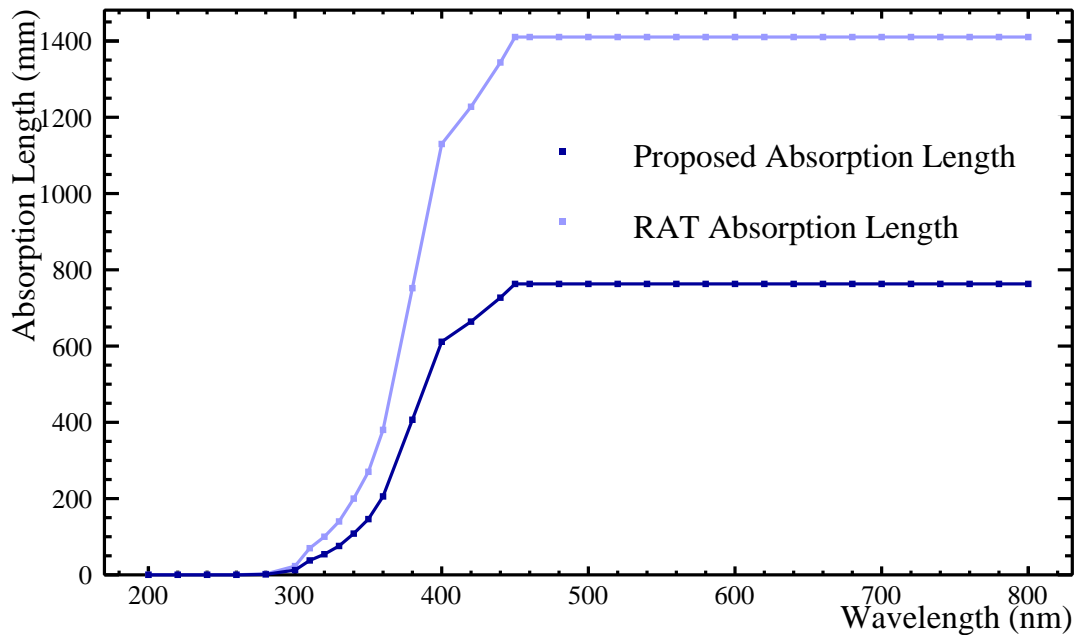
Figure 3.12: Two absorption spectra of acrylic: in the light blue is the initial spectrum from RATDB, in the dark blue is the spectrum proposed after the laserball analysis.

have damaged the surface of the AV. Scratches in the AV surface would lead to an increase in surface scattering. Clearly, surface scattering is not the same as bulk Rayleigh scattering and its wavelength dependence and angular distribution are different. However, it was thought that in the detector simulation, as the AV thickness is small compared to the total length of the light paths of photons, increasing bulk Rayleigh scattering could be a viable alternative to introducing a new model for potential surface scattering [76]. The proposed absorption length for acrylic, after the increase in scattering is taken into account, is shown in Fig. 3.12.

The proposed set of optical constants (the combination of the changes in both waters and acrylic) was tested by use in a SMELLIE analysis. The proposed scattering length of acrylic, in particular, was actually shown to be unphysical by this test. This SMELLIE analysis and the conclusions drawn from it are discussed in detail in Chapter 7. Due to these results, a further updated set of optical constants were calculated.

3.2.3 Updated Optics

After the SMELLIE analysis with the proposed optics constants, the decision was made to revert the scattering length of the acrylic in RATDB to its initial state,
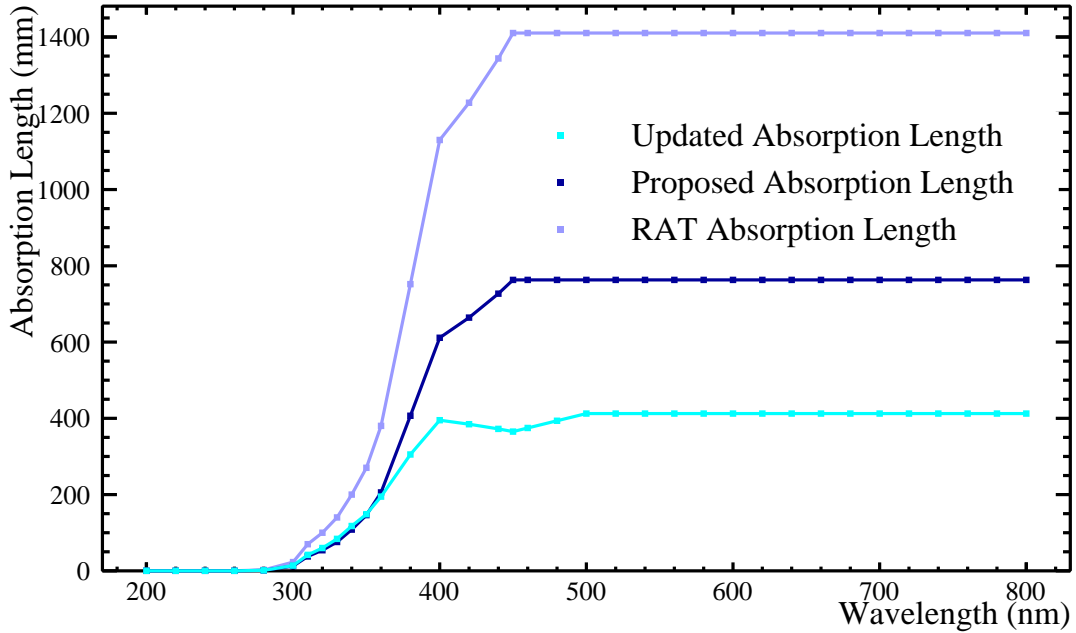
Figure 3.13: Three absorption spectra of acrylic: in light blue is the spectrum from the RAT database, in dark blue is the spectrum proposed after the laserball analysis and in cyan is the updated spectrum following the SMELLIE analysis with the proposed optics constants.

shown in Section 3.2.1. However, the overall attenuation length of the acrylic was still required to decrease, according to the laserball analysis. Therefore, all of extra attenuation has been attributed to absorption. This is shown in Fig. 3.13.

As this updated set of optical constants follows a completed SMELLIE analysis on water data, there is a now an internal water scaling factor that has been calculated: 1.28 ± 0.05(stat.) (see Chapter 7 for details of this analysis and result). Although this scaling factor was only found by analysing possible scattering events in the internal water, as discussed in Section 3.2.2, the decision was made to treat the internal water and external water as optically identical. Therefore, the proposed measured Rayleigh scattering scaling factor is applied to the external water scattering length as well as the internal water.

When the scaling factor is altered in simulation, the absorption length should be scaled in order to keep the attenuation length constant, as the attenuation length is the variable constrained by the laserball measurement. This is done using the relationship:

$$\alpha_{att}^{prompt} = \alpha_{abs} + k\alpha_{RS}, \tag{3.2}$$

where α_{att}^{prompt} is the attenuation coefficient (the inverse of the attenuation length) measured from the prompt light peak in laserball data, α_{abs} is the absorption coefficient and α_{RS} is the Rayleigh scattering coefficient. k is a factor that would be equal to one in an ideal world. The attenuation length in the laserball analysis is measured using the prompt light peak, i.e. direct hits, meaning that scattered light is usually removed from consideration. However, small angles scatters cause scattered photons to still be included in the prompt peak. This lengthens the measured attenuation length and means that, in this case, k has to be less than one [76]. During SNO, k was determined to be 0.820 ± 0.003 [93].

To summarise, the updated set of optical constants includes the initial RATDB scattering length for acrylic, the updated absorption length for acrylic (accounting for the additional attenuation) and the proposed, identical, attenuation lengths for both waters, constrained by the laserball analysis, with the split of the absorption and scattering components determined by the SMELLIE scattering analysis. This set of optical constants is determined by the SNO+ collaboration, via various tests, such as iterative running of the optical calibration analysis and then energy calibration analysis, to be a good measurement of the true optical state of the SNO+ detector during water phase. Therefore the optical constants in RATDB are updated to match and are used by the collaboration henceforth.

Chapter 4

SMELLIE

As discussed in Chapter 2, SMELLIE is an optical calibration system designed to measure the scattering characteristics of the detector media. An overview of the system will be presented here, along with more detail in Section 4.1 about components of the system that are crucial for the work presented in later chapters. For a more comprehensive description of SMELLIE, refer to [53, 94].

The basic principle of the system is shown in Fig. 4.1. There are five injection points and at each of these, three optical fibres are attached to the PSUP. A 5×15 fibre switch connects these fibres to four fixed wavelength lasers. Each of the lasers has a variable attenuator attached enabling tuning of the average of shot energies which can be adjusted further by electronic tuning. Also connected is a supercontinuum laser capable of producing a 10 nm optical-bandwidth beam in the region 400 – 700 nm.

The locations of the nodes of the injection points are spread across the detector. At each node, the three fibres have different pointing angles with respect to the centre of the detector: $0°$, $10°$ and $-20°$. This means a wide range of path lengths though the different areas of the detector are available for analysis.

Not located on the detector, SMELLIE also includes a monitoring PMT unit (MPU). This is formed of a Hamamatsu H10721-210 PMT which receives a fraction of the light from each pulse (via a beam-splitter). The maximum voltage and integral of this pulse are proportional to the energy of the shot (or number of photons in the shot). The MPU therefore enables a shot by shot measurement of the shot energy of SMELLIE pulses.

As previously discussed in Chapter 2, SMELLIE uses the external asynchronous (EXTA) trigger, tuned to a suitable delay. This ensures that the light pulse is recorded by the detector at the same time in each event window, which is crucial for analysis as the time of PMT hits compared to the emission time of the light is used to determine
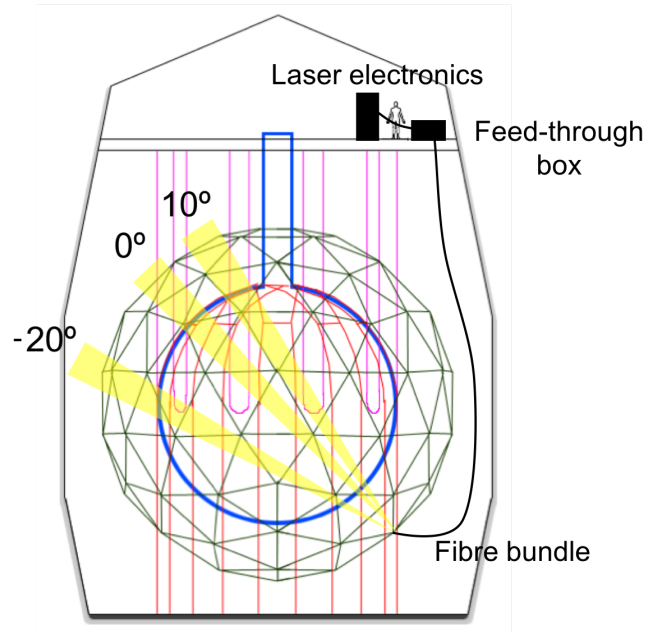
Figure 4.1: A diagram showing the on-detector elements of a single injection node of SMELLIE. Fibres are attached to the PSUP and connected to lasers located on deck. There are 5 injection points and at each of these are three fibres, one pointing at 0° with respect to the PSUP centre, one at 10° and one at -20°. Figure adapted from [53].

which photons are likely to have been scattered. See Section 4.2 for more details on this.

For analysis purposes, each SMELLIE run consists of several subruns. These subruns are identifiable in analysis through the use of an index called the subrun number. Every pulse in each subrun will have identical settings. However, between subruns, hardware settings will change. A record of the planned hardware settings is kept in the SMELLIE CouchDB database as run plans, as well as a set of run description files which records evidence of what was done. One of the purposes of keeping these records is so that subruns can be simulated accurately. Section 4.3 discusses this further.

One of the crucial measurements of the scattering characteristics that SMELLIE enables is of the scattering length. This makes it possible to separate the attenuation length measured by the laserball into scattering and absorption components.

The scattering length calibration is crucial for several reasons, however. It is an input to reconstruction; for example, if the scattering length is shorter than expected, photons will scatter more often in the detector and are therefore less likely to hit a PMT within the event time window, causing the energy reconstruction to be biased towards lower energy [53]. It also affects position resolution, as position is

reconstructed from the position and timing of hit PMTs. Alternatively, a longer scattering length would result in more absorption occurring, for a given total attenuation length, lowering the NHits/MeV. In short, measuring the scattering length over the wavelength range of interest is crucial to understanding the detector response.

Additionally, the scattering calibration is used to tune RAT and therefore produce simulations which accurately model data. It also provides an insight into whether there is any contamination or degradation of the detector medium. For example, contamination in the medium might introduce scattering which is not Rayleigh in nature [87]. Equally, a change in the scattering length over time would indicate that something was changing in the medium (a dramatic example could be Te crystallising out of the complex) and such measurements would be the catalyst for further investigation.

Te-loaded LABPPO is a novel cocktail, developed by the SNO+ collaboration, as discussed in Section 2.2. Therefore its scattering properties are not well studied, making a calibration of them vital. Measurements of the scattering properties of the loaded scintillator have mainly been made on very small samples, such as 1×1 cm cuvettes filled with the cocktail [87]. These measurements indicate scattering consistent with Rayleigh scattering and a scattering length on the order of 10 m [95, 96]. However, the scattering length is extremely difficult to measure as the samples are of a size orders of magnitude smaller than the scattering length and this is compounded at short wavelengths as the scattering length is longer than the absorption length. Therefore, SMELLIE provides critical information by being able to measure the scattering length of Te-loaded LABPPO in-situ.

This work, however, is focused on the commissioning of SMELLIE and the scattering analysis in the water phase of SNO+. The SMELLIE water data set was largely collected by the author and is discussed further in Section 4.3.

Unlike for Te-loaded LABPPO, the scattering behaviour of UPW has been previously measured and is well understood [97]. However, the SMELLIE water data set is of particular interest for several reasons, such as confirmation of the scattering analysis method on a well understood medium. The main reason that this data is crucial is that the scattering and absorption lengths in UPW are long in comparison to the size of the detector, meaning that the data has been taken in a regime of single or no scattering in the detector medium. This allows for the measurement of angular intensity distributions of the laser beams (know as the beam profiles) using detector data (see Chapter 6). This would not be possible in scintillator phase due to the

shorter scattering length and presence of re-emission, which mimics the scattering signal and will be further discussed in Chapter 8.

Before the analysis of SMELLIE data is discussed, the hardware must be understood in greater depth. Therefore, the following section will discuss the major components of the SMELLIE calibration system.

4.1 Hardware

Fig. 4.2 shows a diagram of the complete SMELLIE calibration system with each of the components and the connections between them labelled, including the SNO+ detector systems critical to the functioning of SMELLIE. SMELLIE hardware has been discussed in great detail in other material and therefore only the components of which an understanding is required for this work will each be covered in further detail in this section. For further information on the components not covered here or only discussed in minimal detail, please refer to [53, 94].

Apart from the detector, the components that are physically located within it and the SNO+ systems enclosed in the black border in Fig. 4.2, all the components of SMELLIE are housed in two electronics racks in the DCR.

4.1.1 Lasers

The most integral components of SMELLIE are the five lasers. While the two other ELLIE subsystems use LEDs as their light sources, lasers were chosen for SMELLIE for several reasons. The emitted light is coherent, meaning that it allows for a higher shot energy and more effective collimation. Collimation, particularly, is important for SMELLIE as the scattering analysis relies on the identification of photons which have scattered out of the beam. This will be discussed further in Chapter 7. In addition, the emitted light has a narrow spectral bandwidth which is useful for studying scattering characteristics as a function of wavelengths. Using lasers also allows shorter pulses to be generated than using LEDs and these pulses remain shorter when travelling through the fibres because the narrow bandwidth means they experience little chromatic dispersion.

The five lasers consist of four pulsed-diode laserheads and one supercontinuum laser. Both of these will be discussed in the following two sections.
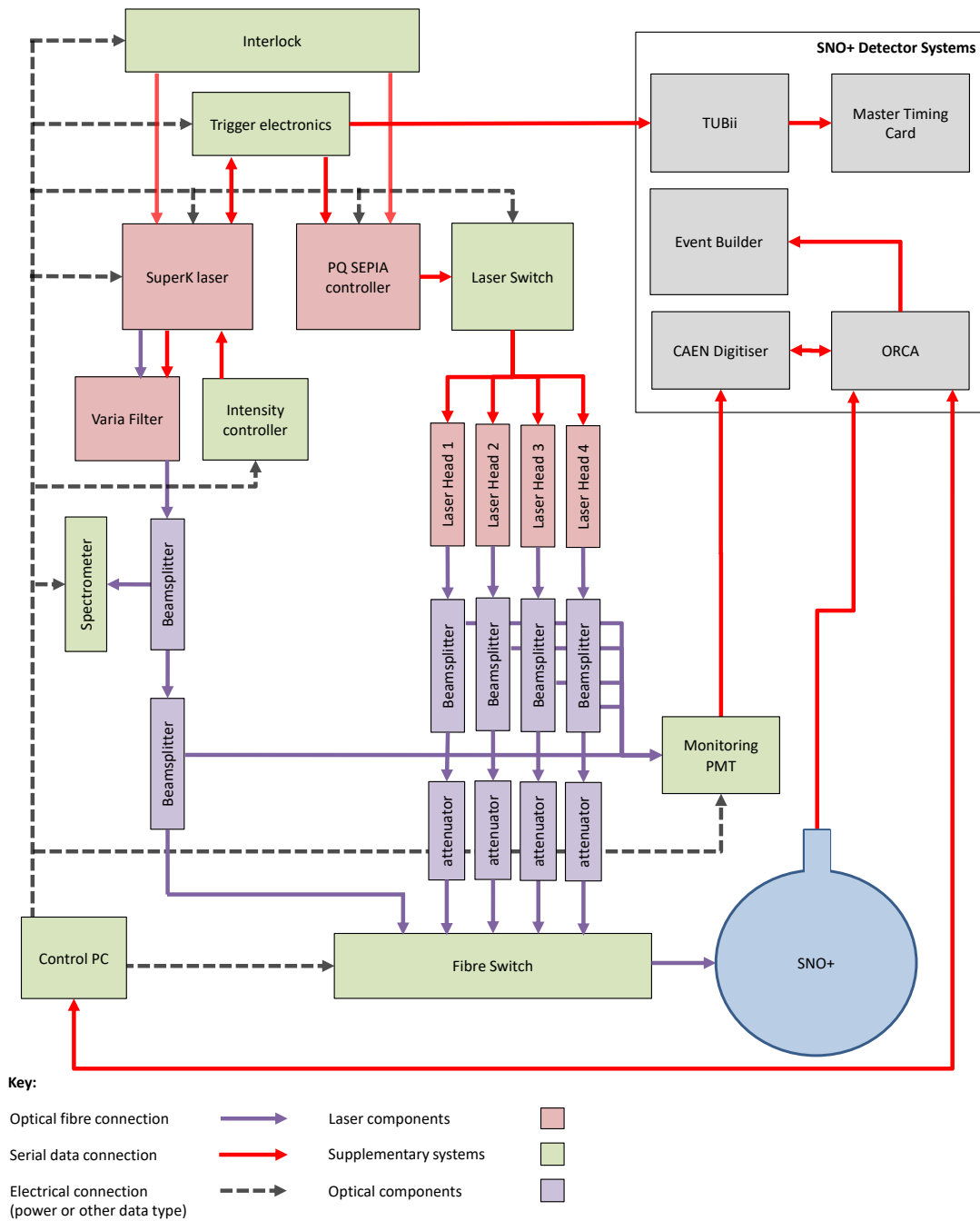
Figure 4.2: A diagram showing the on-detector elements of a single injection node of SMELLIE. Provided by J. Lidgard and adapted from [98].
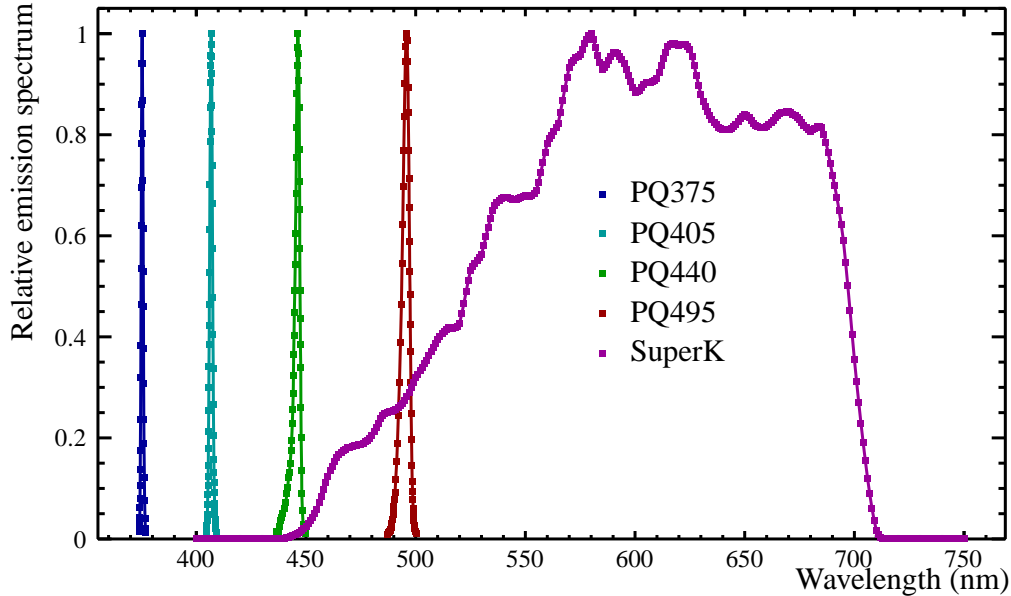
Figure 4.3: Relative emission spectrum versus wavelength for each of the five lasers, as recorded in RATDB. The spectra are normalised such that they peak at 1. For the supercontinuum laser, 10 nm bandwidths of this wavelength spectrum are selected when running a simulation.

PQ Lasers

The four diode laserheads used in SMELLIE are manufactured by PicoQuant and are therefore often referred to in this work as PicoQuant (or PQ) lasers [99]. These nominally operate at wavelengths of 375, 405, 440 and 495 nm and were chosen due to the range of high efficiency of the PMTs (shown in Section 2.4). The wavelengths were also selected to avoid high absorption wavelength regions of LAB loaded with ^{150}Nd. Clearly, ^{150}Nd is not being used in SNO+; however, it was the original proposal for the $0\nu\beta\beta$ isotope for SNO+. Fig. 4.3 shows the relative emission spectra for each of the lasers (including the supercontinuum laser, discussed in the following section) as used in RAT simulations. These wavelengths are taken from the laser documentation provided by the manufacturer [94].

The lasers also each have relative emission time spectra specified in RATDB, taken from manufacturer documentation [99] and shown in Fig. 4.4.

Each PQ laser head is connected to an attenuator. These are manufactured by Fibredyne and are manually adjustable [94]. Therefore, any adjustments require an operator underground making this a rare procedure. It allows the useful range of shot energy settings of the laser to be shifted, such that the laser can be operated at a higher shot energy while producing the same amount of light in the detector. One
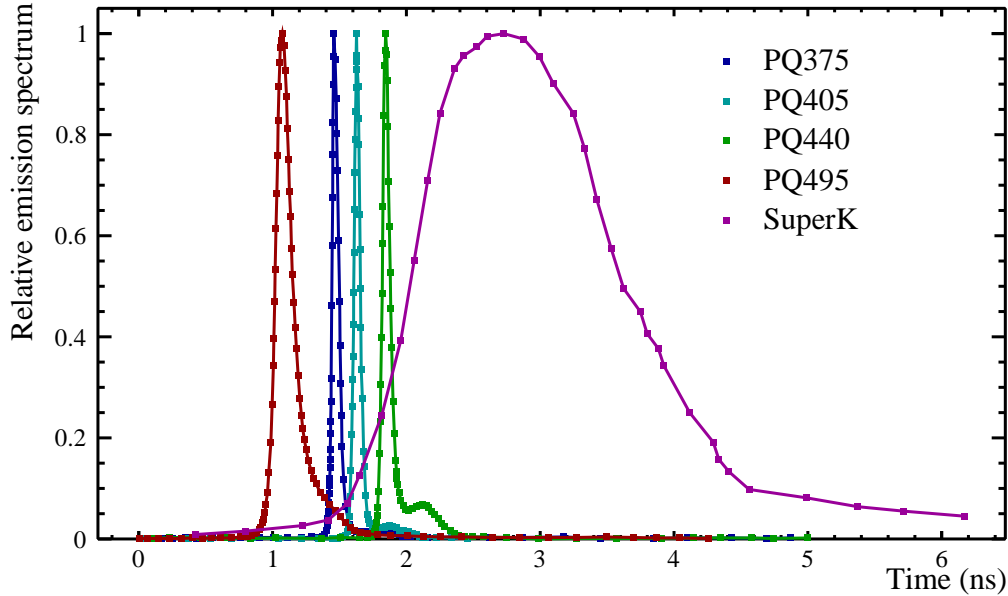
Figure 4.4: Relative emission spectrum versus time for each of the five lasers, as recorded in RATDB. The spectra are normalised such that they peak at 1.

reason to do this is discussed in Section 5.5.

The PQ lasers are controlled and powered via the SEPIA II laser driver [100], also manufactured by PicoQuant. This allows control of the pulse energy and the repetition rate. The pulse energy can be set as a percentage of the maximum drive pulse energy in increments of 0.1 % from 0 – 100 % [94]. In practice, the shot energy settings for SMELLIE are only adjusted in 1 % increments and only the ranges between approximately 80 – 100 % provide useful operational settings. The optical output pulse energy from the lasers is a very non-linear function of the drive pulse energy.

While the SEPIA does provide preset triggering rates in order to easily control the repetition rate, all the these are too high for use in SMELLIE [53]. Instead, the repetition rate is controlled through the SEPIA by using an external trigger. This takes the form of a voltage pulse generated by the MPU (discussed in Section 4.1.4) [94], allowing the PQ lasers to trigger at frequencies more desirable for SMELLIE, typically 1 kHz.

The SEPIA unit used in SMELLIE contains a single laser driver, a decision driven by investment costs and the fact that only one PQ laser is fired during each subrun. However, this means that a mechanism is required to physically change which laser head is connected to the SEPIA when a change in wavelength is requested, as all

four PQ lasers cannot be constantly connected. Clearly, this cannot be done by an operator every time a SMELLIE run is taken. To solve this problem, the laser switch unit was manufactured at the University of Oxford by the Central Electronics Group [94]. This provides an automated method of switching which laser head is connected to the SEPIA.

No light is routed through the the laser switch. Rather, electronic signals from the SEPIA are passed to the laser switch instead of the laser head. There are six output sockets on the laser switch and the electronic signal is routed to a selected one. Each socket is connected to a laser head (two are spare). This routing is achieved through electromechanial relays, meaning that only one output channel can receive the signal from the SEPIA at a given time. This switching process does take a finite amount of time (between 8 and 9 s), particularly as the SEPIA is powered down every time a switch is performed, and the laser switching process is a limiting factor in the length of time between SMELLIE subruns.

SuperK Laser

The fifth SMELLIE laser is a SuperK Compact laser (henceforth called the superK), manufactured by NKT Photonics. This is a white light source, consisting of a super-continuum laser capable of producing a spectrum of light from 450 – 2400 nm, with a pulse rate of 1 – 20000 Hz [101].

A central wavelength between 400 - 700 nm can be selected, with a beam width as narrow as 10 nm. This is achieved using the SuperK Varia (a filter box using a variable bandpass filter), also produced by NKT Photonics [102]. The available spectrum of the superK laser, as measured by J. Lidgard and as in RATDB, is shown, alongside the PQ lasers, in Fig. 4.3.

Similarly, in Fig. 4.4 is the timing spectrum of the superK, again, as implemented in RATDB and compared against the PQ lasers. It is clear that the superK has a much larger intrinsic pulse time FWHM than the PQ lasers of approximately 2 ns [101].

The superK laser, as supplied, comes with an inbuilt neutral density (ND) filter; the operation of which has been disabled by the manufacturer for fear of degradation of the filter from high average power operation. However, J. Lidgard enabled the use of this ND filter through an external Arduino controller. As the laser is only operated at very low average powers there is no damage to the filter. This allows the filter to be, in essence, a variable attenuator for the superK. Due to this filter, the light from the superK is not passed through an additional attenuator, as the beams from the PQ lasers are [103].

4.1.2 Optical Fibres

Optical fibres are a large component of the SMELLIE system. They are used for connecting the various components together as well as transporting the light into the detector and as such are split into two corresponding groups - rack fibres and detector fibres.

There are 15 detector fibres, each 45 m long. These are the fibres that begin at the rack (specifically at the fibre switch, discussed in Section 4.1.3) and descend into the SNO+ cavity, each one ending by being mounted on a PSUP node, from which light is emitted through the detector. At the detector end the fibres terminate in collimators, discussed shortly. At the rack end, the fibres are FC/APC terminated, which means that they end in a angle-polished ferrule connector [94].

For clarity, each detector fibre has been named for the node they are attached to on the PSUP. These are 07, 37, 25, 55 and 21. However, the fibres on the 21 node are referred to by number 93 due to mislabelling on the end of the fibres in the detector as they were originally spares. However, interest was shown in having fibres mounted at the top of the detector, next to the neck, pointing almost vertically downwards and illuminating the bottom of the AV where they might be used to detect any solids that might precipitate from the Te-loaded scintillator. Therefore the installation of these fibres was planned after the detector fibres had already been lowered into the cavity and filling of the AV had begun. Thus fibres that had originally been intended as spares were mounted at the 21 node. As this node is close to the neck, the beams emitted from this point are particularly obscured, both by the neck itself and by the hold-down ropes which keep the AV from rising due to its buoyancy once filled with liquid scintillator. This is discussed further in Chapter 6. Despite this, the beams are of potential interest in the detection of vertical variations of optical properties or precipitation.

The other four nodes were chosen to provide a variety of path lengths through the detection medium and through all parts of the detector. Fig. 4.5 shows the five injection points of the SMELLIE fibres and the direction of the 15 beams emitted from them overlain on a sphere representing the AV. The injection position is often referred to as the fibre position as it is the location of the tip of the fibre. Likewise, the beam direction is generally called the fibre direction as this is the pointing direction of the fibre tip.

As there are three fibres at each node, the node number in the fibre name is preceded by 0, 1 or 2, as shown in the legend of Fig 4.5. This number was originally meant to correspond to the mounting angle of each fibre: $0°$, $10°$ or $-20°$ between the direction
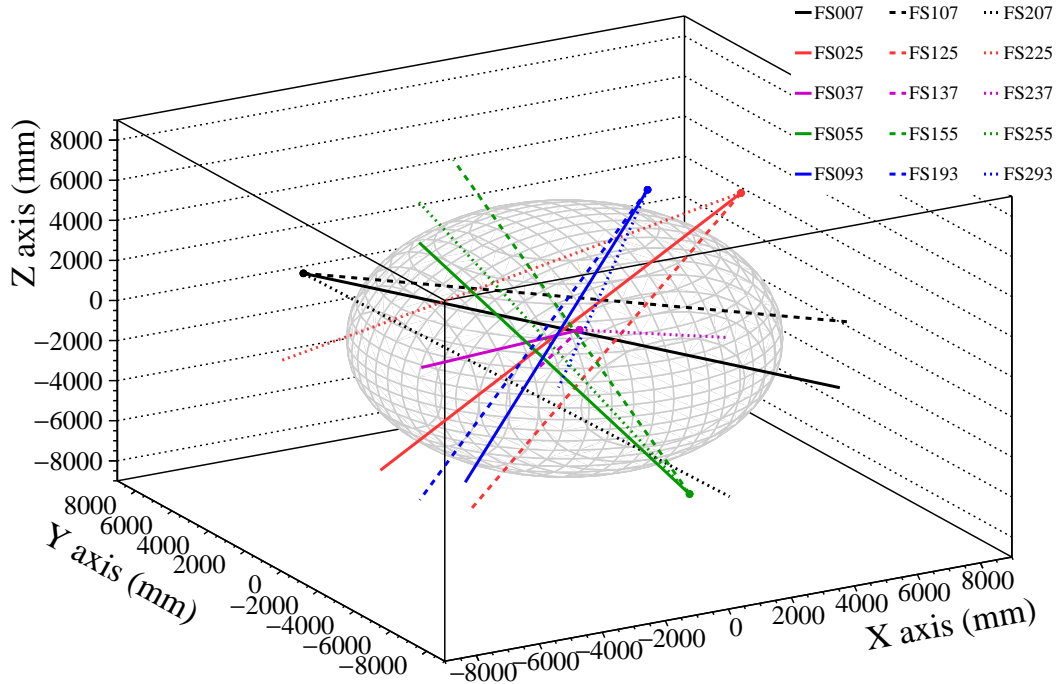
Figure 4.5: The mounting points and beam directions for all 15 SMELLIE fibres, overlain on a spherical representation of the AV. Adapted from E. Leming.

of the beam and centre of the detector. However, due to installation complications, this correspondence is not reliable. Table 4.1 lists each of the SMELLIE fibres along with their pointing angles.

There are many more than 15 rack fibres, as these fibres connect all of the SMELLIE components that require an optical connection and are housed in the two racks. These fibres are all 1 m in length and both ends are FC/APC terminated.

All the fibres used in SMELLIE come from the Corning InfiniCor SXi 50 μm multimode range [94]. These fibres have very low intrinsic radioactivity. They are mounted on the PSUP alongside the PMTs. Therefore, the products of any radioactive decay would be able to enter the PMTs with barely any shielding. To alleviate this concern, it was calculated that for all the SMELLIE fibres, the total radioactivity is lower than that arising from the PMTs themselves [53].

These fibres are graded index fibres, in which the refractive index of the core reduces with distance from the centre of the fibre core. This means that there is low modal dispersion within the fibres and hence sharp emission time profiles. They are multimode; however, the optimal operating wavelength range for these fibres is between 750 – 1450 nm (specifically, they are OM3 telecoms style fibres). As already discussed, SMELLIE uses wavelengths between 375 - 700 nm. This means that extra

Fibre Name	Node	Pointing Angle (°)
FS007	07	0
FS107	07	10
FS207	07	20
FS037	37	10
FS137	37	0
FS237	37	20
FS025	25	0
FS125	25	10
FS225	25	20
FS055	55	10
FS155	55	20
FS255	55	0
FS093	21	0
FS193	21	10
FS293	21	20

Table 4.1: The 15 SMELLIE fibres listed alongside their injection nodes and pointing angles.

attenuation is caused by the optical attenuation through the fibres in SMELLIE. For a further discussion of this, see [53].

As the name implies, the core diameter of the fibres is 50 μm. This is surrounded by glass cladding and a plastic sheath. This design and the refractive indices of the materials involved mean that the beam emitted from the fibres is more divergent than desired for SMELLIE [53], hence the requirement for collimators on the detector end of the fibres.

Collimators

As just discussed, the detector fibres are collimated at the detector end for the purpose of the scattering analysis. The collimators used were designed by K. Clark and primarily manufactured at the University of Oxford (except from the glass ferrules, which were manufactured by ThorLabs) [94].

The crucial component of each collimator is a grin lens. This is a cylindrical quartz rod in which the refractive index varies such that it is less at the edges than in the centre. This has the effect of collimating the light passing through the rod, with the focal length dependent on the length of the rod.

The collimators were designed to reduce the half opening angle of the beam to $\sim 3.5°$ [53]. In practice, when Gaussians are fitted to the central regions of the beams from detector data, half opening angles from $0.7° - 5.0°$ are observed. This variation

is due to speckle patterns in the central region of the beam alongside production variations in the collimators themselves. The beams also have significant tails, arising from reflections and scattering inside the collimators. The beam profiles are discussed at length in Chapter 6.

Mounting

The mounting of the fibres onto the PSUP has already been discussed in Section 2.6.4 and a diagram of the mounting plates used for all the SMELLIE systems is shown, attached to a PMT bucket, in Fig. 2.10. These plates are fixed to the buckets using plastic rivets; as already discussed, low radioactivity is crucial in all the components located in the detector. The plates are made from Delrin, which, alongside its intrisic low radioactivity, resists deformation in long term water exposure. This is also critical in components which will be submerged with the detector for approximately 10 years [94].

Once the installation of the 15 detector fibres had occurred, verification of the mounting positions and directions was conducted.

Neither verification is fully covered in this work. The author undertook basic checks to compare the positions of the beamspots for each fibre in simulation and data. This gives confidence that the mounting positions were recorded correctly by the installation team, backed up by investigations conducted by S. Langrock [104].

The fibre directions have more freedom than the fibre positions. In addition to small movements of the collimators themselves within the mounting plate, some fibres were installed at an angular position not corresponding to their name. Recognising if a fibre is installed at a different angle with respect to the detector centre than expected is simply achieved by the comparison of data and simulation. Therefore, the approximate pointing angle of each fibre is known.

The very first SMELLIE runs taken with a UPW filled detector at HV were used to refine the fibre direction calculations. This was done by T. Tunstall [105] by fitting Gaussians to the observed NHits in the region of the expected beamspot for each fibre and then calculating the vector between the assumed fibre position and the peak of the Gaussian. The directions were then updated in the SMELLIE database in RATDB. Nonetheless, when the updated directions are used to simulate SMELLIE the match between the positions of the beamspots in data and simulation is visibly not perfect for some fibres. This is not a concern for this thesis, due to the work

discussed in Chapter 6. Furthermore, the systematic error arising from realistic mis-modelling of the fibre direction on the calculation of the scattering length is calculated in Section 7.4.8.

Beamsplitters

As well as optical fibres, beamsplitters are present along the optical paths of SMEL-LIE. These are used to split the light - for example, to send a proportion of each pulse to the MPU (discussed in Section 4.1.4) and the remainder to the detector. They are of fused fibre taper design and are manufactured by Wealth & Grace Alliance Ltd.. For further details, such as the splitting ratios, refer to [53].

4.1.3 Fibre Switch

The fibre switch is the last SMELLIE component that light passes through before entering the detector fibres. This is manufactured by Leoni and distributed by Laser Components (UK) Ltd. [53].

SMELLIE requires that light from any one of the five lasers can be emitted from any one of the 15 detector fibres. Therefore, a switching mechanism is required such that the output from any of the lasers can be routed to the desired fibre. This is the role of the fibre switch.

The fibre switch in SMELLIE actually consists of two fibre switches. This is, again, a consequence of the 21 node fibres being installed at a later date. The original fibre switch unit is a five input to 14 output unit. Clearly, for 12 detector fibres, 14 outputs are sufficient (two outputs were in reserve). However, for 15 detector fibres, further outputs are required. Therefore, a second fibre switch unit was installed. This is a smaller unit with one input and four outputs. The two units are connected by an output on the larger one leading to the input of the smaller unit. The 12 original detector fibres and FS093 are connected to the other 13 outputs of the larger unit, with FS193 and FS293 connected to the outputs of the smaller unit. Therefore, the light pulses emitted from FS193 and FS293 are subject to higher levels of attenuation than those from the other detector fibres as they pass through one more fibre switch. This extra attenuation is one of the many variables to calibrate during the process of commissioning SMELLIE, such that it can be adjusted for by running these fibres at higher shot energy settings. This is discussed in Chapter 5. Intensity loss in the fibre switches is caused by a variety of factors including fibre misalignment, switching block misalignment and intrinsic losses in the internal fibres [53].

Like the laser switch, the switching mechanism in the fibre switch does take time, between 1 ms and 1 s. However, as no hardware is powered down for each switch, the amount of time required is smaller than for the laser switch. In practice, it means that the same laser is run through every fibre before switching lasers, rather than vice versa.

For details on the control and mechanism of the fibre switch, refer to [53].

4.1.4   Monitoring PMT Unit

SMELLIE requires some method of in-situ monitoring as the shot energy of a laser can vary from pulse to pulse. This variation is dependent on the intensities that the laser is operating at - at low shot energies, close to the lasing threshold, the variation is larger than at higher shot energies. SMELLIE is operated at a range of shot energies, making a method of monitoring this instability all the more critical. Having a monitoring system allows the user to determine if variations in the number of hit PMTs during a subrun is due to changing detector conditions or due to the shot energy of the laser fluctuating. It can also point towards abnormal laser behaviour if large shot energy variations are present in high shot energy regimes.

The way this shot by shot monitoring is achieved for SMELLIE is via a monitoring PMT unit (MPU). This consists of an ex-situ PMT, a Hamamatsu H10721-210, with a 10 mm diameter face. This is housed in a die-cast aluminium box, designed and built by J. Saunders at the University of Oxford. The PMT is sensitive to wavelengths between 200 – 700 nm, fully covering the entire wavelength range of SMELLIE, with the peak efficiency at approximately 400 nm. A crucial reason for the selection of this PMT is that the rise time is 0.57 ns, meaning that every laser pulse (even at 2kHz) will be detected individually, cleanly and with high time resolution [94].

The output pulse from the PMT in the MPU is shaped and slowed, making analysis of the pulse easier. This output pulse is the sent, via BNC, to the CAEN-ADC, discussed in Section 2.5. From here, it is accessible to the operator and the digitised output waveform of the MPU is recorded for each event alongside the other detector data.

The peak height of the output PMT pulse is correlated to the shot energy of the laser pulse, as originally investigated by K. Majumdar [53]. This means that the peak height of the MPU CAEN trace is a variable used in SMELLIE analysis to monitor the pulse by pulse shot energy of the lasers. As well as monitoring laser stability, this also makes an analysis correction to any pulse energy variation possible.

The gain of the MPU PMT can be controlled by the operator. For details of how this is achieved in hardware, please refer to [53]. This is necessary as the CAEN-ADC only accepts signals up the 2 V. Therefore, information from any MPU signal that exceeds this voltage from the MPU would be lost and the pulse height would not be calculable. Optimising the gain to be suitable for the laser/wavelength/shot energy combination used in any given subrun is a commissioning activity discussed in Section 5.6.

4.2 Trigger

As previously discussed in Chapter 2, there are both master and slave modes used for triggering with the EXTA trigger in optical calibration runs. Very early SMELLIE testing was performed in slave mode (as the PQ lasers can be run in either mode), before this work began. However, the superK laser does not support slave mode. Therefore, since its installation in 2016, SMELLIE is run purely in master mode.

For the PQ lasers in master mode, the trigger signal is a TTL voltage pulse produced by the National Instruments DAQ Unit (henceforth called the NI Unit), located in the SMELLIE racks. This signal is split, with one half converted to an ECL pulse and sent to TUBii where it is delayed by a controllable delay and then sent on to the MTC/D, where a global detector trigger (an EXTA trigger) is issued. The other half is sent to SEPIA, in order to trigger the PQ lasers to fire.

However, for the superK, due to the light generation mechanism, the timing of the light pulse has a large uncertainty relative to the electrical pulse triggering the light; this optical pulse to electrical pulse jitter has a magnitude of up to 2 μs [101]. This means that to accurately trigger the detector such that the light arrives at the same time in each event window, the trigger must arise from the emissing of the light pulse itself. To this end, the superK has an inbuilt photodiode on its output which generates a pulse when it senses light. This pulse is then converted to ECL format and sent to TUBii in order for a detector trigger to be issued, as for the PQ lasers.

In both these cases, in order to ensure that the light pulses arrive in the detector in the middle of the event window, a delay is applied to the signal by TUBii. For the PQ lasers this is 230 ns and for the superK 200 ns. This arrival of light at the same time with respect to the trigger in every event is explored in detail in Section 5.3 as a commissioning activity.

4.3 SMELLIE Runs Settings

Details of the SMELLIE runs taken during water phase and used in this work are discussed in the next chapter, in Section 5.7. This section, however, discusses the general principles of SMELLIE runs, such as the trigger settings.

During a SMELLIE run, the EXTA trigger is used for SMELLIE pulses. The PulseGT (referred to as PGT) is also kept active. This is a forced detector wide trigger which is issued periodically (at a rate of 50 Hz) and used to measure the noise rate in the detector. This information is used in the analysis to simulate realistic noise levels and to correct for noise in the commissioning and thus infer how much light in the detector is truly caused by SMELLIE pulses. These are the only two triggers activate during the SMELLIE runs used for this analysis.

As discussed, SMELLIE runs consist of many different subruns, each with a different combination of SMELLIE hardware settings: for example, different lasers, shot energies, wavelengths, frequencies, gains or fibres. Due to the large number of SMELLIE settings therefore required for each subrun, a run plan file is programmed in advance which includes all of the SMELLIE settings for the entire run. This is simply loaded into ORCA, the detector control software, by an operator. ORCA communicates with a server program running on the SMELLIE control PC (Snodrop), which then controls the SMELLIE hardware to achieve the settings described in the run plan. This provides more reliable SMELLIE control as the control of SMELLIE hardware is abstracted out of ORCA, meaning that safety procedures such as interlocks and laser shut-downs can be operated even if ORCA crashes. The run plan, once loaded into ORCA, can be run without user intervention and without a SMELLIE expert present.

Each pulse within a subrun has identical settings hardware settings. The run description file represents a log of what actually happened during a SMELLIE run. This is not necessarily the same as the run plan and can be shorter because a run can be prematurely terminated due to user intervention or due to a hardware error. This run description file can therefore be used to identify the settings used for each subrun in analysis. The subruns are identifiable in analysis by the use of a subrun number variable.

Additionally, there are hardware configuration files stored in the SMELLIE CouchDB database. These describe the routing of the fibres into and out of the fibre switch and laser switch, allowing Snodrop to know which switch channel to use for a desired detector fibre or which laser switch port to activate for the chosen laser.

Chapter 5

SMELLIE Commissioning

Prior to this work, SMELLIE had only been operated in the detector a handful of times. These test runs had taken place when the detector was partially filled with water, with only half the fibres installed (those on the upper half of the PSUP were installed in 2016) and with only a fraction of the PMT crates at HV. Therefore, these runs have limited use for testing some basic SMELLIE functionality.

Consequently, a large portion of this work is dedicated to the initial commissioning of SMELLIE in the water phase of SNO+. This includes checking basic properties of the system, resolving discrepancies with simulation, investigating unexpected features and working out how to relate the tuneable variables of the system to observables quantities in the detector, all while ensuring that SMELLIE is never fired at settings which could cause damage to the PMTs or detector electronics.

As the commissioning work was split between the author and J. Lidgard, this chapter will explore three areas of the commissioning that were the work of the author (Sections 5.3 – 5.5) and one area of commissioning that was jointly worked upon (Section 5.6). Many other SMELLIE commissioning activities not discussed here were performed during this work; the selection presented in this chapter should provide an overview of how to progress from a system that has rarely been fired before to one that is regularly used to take reliable data for analysis, as well as many of the challenges encountered during that process. A brief overview of the data used both in this chapter and the next two chapters will also be presented in Section 5.7.

Before the commissioning work can be presented, however, a couple of concepts need to be discussed. One is the coordinate system that will be used henceforth when discussing photons emitted from a SMELLIE fibre (see Section 5.1) and the other is a statistical correction used in the case of multiple photons entering a PMT in a single event (see Section 5.2).

5.1 SMELLIE Coordinate System

When discussing the beams emitted from the SMELLIE fibres, the detector coordinate system is not used. Instead, a coordinate system designed for SMELLIE beams is used. This was partially created by K. Majumdar and S. Langrock in [53, 104]. However, it has since been adapted to the description of non-rotationally symmetric beams by the author.

The coordinate system is polar, centred around a point named the beamspot centre. This point, in detector coordinates (also called PSUP coordinates, as they are centred on the PSUP centre), is found by propagating the fibre direction through the detector from its injection point to the opposite side of the PSUP. This calculation relies on spherical geometry - effects from the AV within the PSUP are completely ignored. (The effect of this simplification is addressed in Chapter 6.) Take the equation for a sphere:

$$\left\| \vec{b_p} - \vec{c} \right\|^2 = r, \tag{5.1}$$

and for a straight line:

$$\vec{b_p} = \vec{f_p} + d\vec{f_d}, \tag{5.2}$$

and substitute $\vec{b_p}$ in the former equation with the latter. In these equations, $\vec{b_p}$, is the point being found, i.e. the intersection of the sphere and the straight line (a.k.a. the beamspot centre). $\vec{c}$ is the centre of the sphere and r is its radius. $\vec{f_p}$ is a point on the line (in this case, the fibre position), $\vec{f_d}$ is the unit direction vector of the line (i.e. the fibre direction vector) and d is the distance between $\vec{b_p}$ and $\vec{f_p}$ (along the line).

This results in

$$d = \frac{-(\vec{f_d} \cdot (\vec{f_p} - \vec{c})) \pm \sqrt{(\vec{f_d} \cdot (\vec{f_p} - \vec{c}))^2 - \left\| \vec{f_d} \right\|^2 (\left\| \vec{f_p} - \vec{c} \right\|^2 - r^2)}}{\left\| \vec{f_d} \right\|^2}. \tag{5.3}$$

This equation can be further simplified as the PSUP coordinate system is being used. Therefore, $\vec{c} = (0, 0, 0)$. Also, $\vec{f_d}$ is a unit vector. This simplifies the above equation to

$$d = -(\vec{f_d} \cdot \vec{f_p}) \pm \sqrt{(\vec{f_d} \cdot \vec{f_p})^2 - (\left\| \vec{f_p} \right\|^2 - r^2)}. \tag{5.4}$$

$\vec{f_p}$ is a point on the sphere. Therefore, $\left\|\vec{f_p}\right\|^2 - r^2 = 0$, leaving $d = -(\vec{f_d} \cdot \vec{f_p}) \pm \sqrt{(\vec{f_d} \cdot \vec{f_p})^2}$. One of the solutions to this is trivially zero, which is the case where the intersection of the line and the sphere is f_p (i.e. the fibre position). Therefore, the solution for d at the beamspot centre is $d = -2(\vec{f_d} \cdot \vec{f_p})$.

This is then substituted back into Eq. 5.2, resulting in the location of the beamspot in PSUP coordinates being $\vec{b_p} = \vec{f_p} + 2(\vec{f_d} \cdot \vec{f_p})\vec{f_d}$. This defines the beamspot centre purely in terms of the fibre position and direction in PSUP coordinates.

As mentioned above, the beamspot centre, in the SMELLIE coordinate system, is the origin. Once this location is defined in detector coordinates, each of the PMTs surrounding the beamspot centre can be represented in the SMELLIE coordinate system by using two angles. The point on each PMT which is represented by these two coordinates is the centre of the PMT face.

The polar angle, α, is the easier of the two angles to calculate. It actually does not require the beamspot centre location to be known, only needing knowledge of the fibre position, direction and the position of the PMT of interest in PSUP coordinates. It is calculated by finding the vector from the fibre position to the PMT position ($\vec{f_t p}$) and then using the inverse of the dot product with the fibre direction, as follows:

$$\alpha = \text{acos}\left(\frac{\vec{f_d} \cdot \vec{f_t p}}{\left\|\vec{f_d}\right\|\left\|\vec{f_t p}\right\|}\right). \tag{5.5}$$

A diagram showing α, along with the vectors required to calculate it, is in Fig. 5.1.

On the other hand, calculating the azimuthal angle, ϕ, does require knowledge of the beamspot centre position. A $\phi = 0$ angle has to be defined for each fibre, which is taken to be the vector maximally pointed upwards. Therefore, a plane is defined which is perpendicular to $\vec{f_d}$. $\phi = 0$ is found by projecting a z-direction unit vector into the plane (resulting in a vector known as $\vec{z_{proj}}$).

For each PMT, the vector from the beamspot centre to the PMT of interest is projected into this plane ($\vec{b_t p_{proj}}$). To find ϕ, the angle between $\vec{z_{proj}}$ and $\vec{b_t p_{proj}}$ needs to be found. The direct angle between the two vectors is simple to find using the dot product. However, the rotation direction must always be the same so that ϕ ranges from $0 - 2\pi$ rad (or $-\pi - \pi$ rad), rather than $0 - \pi$ rad. Therefore, the dot product is not enough: if the direct angle is represented as θ, $\phi = \theta$ or $\phi = 2\pi - \theta$. This is shown in Fig. 5.2.

For the SMELLIE coordinate system, the direction of rotation for ϕ is defined as anti-clockwise.
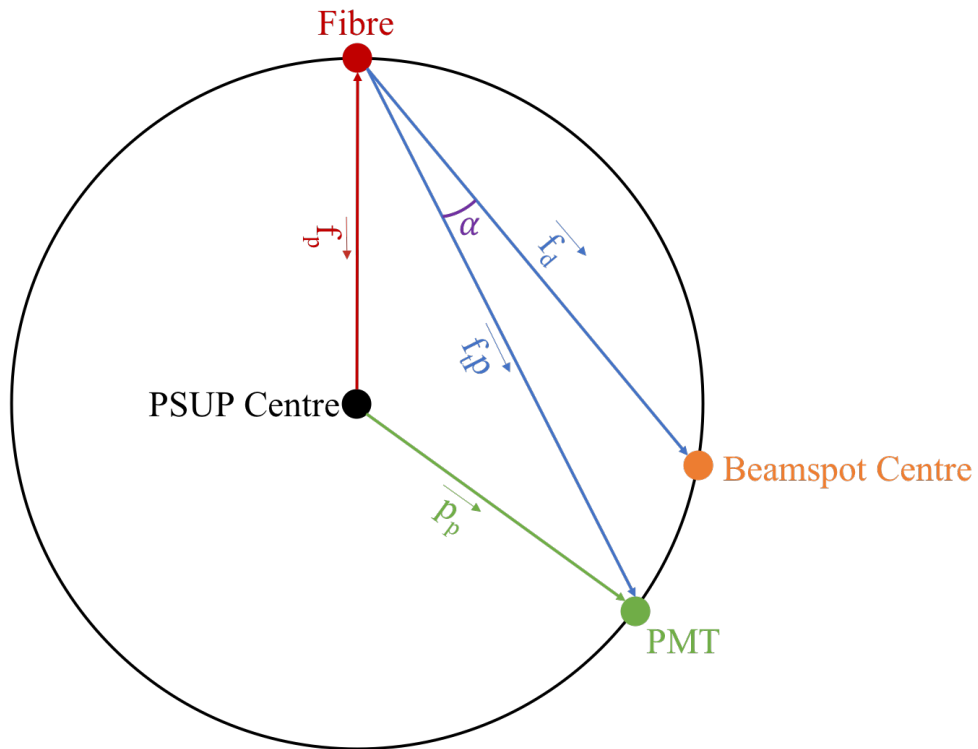
Figure 5.1: Diagram of the definition of the polar angle, α, in the SMELLIE coordinate system, calculated for a fibre which is not pointed through the PSUP centre. The positions of the PSUP centre (black), fibre (red), beamspot centre (orange) and a PMT (green) are shown, alongside their position and direction vectors.
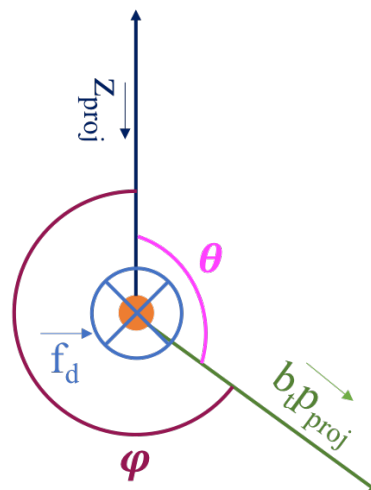
Figure 5.2: Diagram of the definition of azimuthal angle, ϕ, calculated in a plane perpendicular to the fibre direction. The position of the beamspot centre (orange) is shown, alongside the z direction vector and the vector from the beamspot centre to the PMT of interest projected into the same plane. The direct angle between the two vectors, θ is shown, as well.

The dot product of the two projected vectors, as discussed, results in $\cos(\theta)$ or $\cos(\phi)$, multiplied by the magnitude of the two vectors. Meanwhile, the cross product of the same two vectors gives their magnitudes multiplied by $\sin(\theta)$ multiplied by a unit vector whose direction is given by the right-hand rule ($\vec{n}$). The orientation of this unit vector is dependant on whether ϕ is $<$ or $> \pi$. When $\phi > \pi$, $\vec{f_d} \cdot \vec{n} = -1$ and when $\phi < \pi$, $\vec{f_d} \cdot \vec{n} = 1$.

Therefore, $(\vec{b_t p}_{proj} \times \vec{z}_{proj}) \cdot f_d$ results in $\left\| \vec{b_t p}_{proj} \right\| \times \| \vec{z}_{proj} \| \times \sin \phi$.

Consequently,

$$\frac{(\vec{b_t p}_{proj} \times \vec{z}_{proj}) \cdot f_d}{\vec{z}_{proj} \cdot \vec{b_t p}_{proj}} = \tan(\phi). \tag{5.6}$$

This means that ϕ can be found by applying the `arctan2` (from the C++ `math.h` library) function to Eq. 5.6. (`arctan2` is used as this returns a ϕ value between π and $-\pi$, rather than a ϕ value between $\frac{\pi}{2}$ and $-\frac{\pi}{2}$.)

Therefore, a polar coordinate system has been created which defines the location of each PMT relative to the beamspot centre by two angles (α, ϕ) where α is the polar angle and ϕ is the azimuthal angle.

5.2 Multi-hit Correction

As previously discussed Section 2.4, charge is not typically used to discriminate p.e. in SNO+. Therefore, for each PMT hit (excepting noise), it is only known that at least one p.e. caused the hit. If there are multiple p.e., only one hit will be observed. Statistically correcting for this will be referred to as multi-hit correction. This method of statistical correction was developed by J. Dunger [106].

The number of p.e. in a given PMT in a given event is assumed to follow a Poisson distribution, as the probability of creating a p.e. in any given PMT and event is independent of time in the event and independent of other p.e. being created previously in that PMT and event. Therefore, $P^i = 1 - e^{-\mu^i}$, where P^i is the probability of a hit observed in PMT i and μ^i is the expected number of p.e. in that PMT.

In order to find μ^i, P_i must be known. In SMELLIE events, many shots are fired under assumedly identical conditions. Therefore, P^i can be estimated for a specific PMT via its average over many events: $\hat{P} = \frac{N_{Hits}}{N_{Shots}}$. This is known as the occupancy of a PMT. This means for that specific PMT $\hat{\mu} = -\log\left(1 - \hat{P}\right)$.

At an occupancy of one (i.e. a hit is observed in the PMT in every event), it is not possible to know how likely one photon is to create a p.e. in the PMT. Therefore, to

use this method of multi-hit correction, the shot energy of the lasers cannot be too high - this will be discussed more in Section 5.6.

5.3   Time Residual Spectrum

Time residuals have been introduced in Chapter 2. As mentioned there, in the case of SMELLIE, the path calculated by `LightPathCalculator` (and therefore used to calculate the time residual) is the shortest path between the hit PMT and the fibre position.

The use of time residuals is desirable for SMELLIE analysis because time residuals take into account refraction at material changes and the group velocity in materials, meaning that this quantity is more material and fibre position independent than the calibrated hit time [53].

Fig. 5.3 shows the hit time and corresponding time residual spectrum for a subrun of SMELLIE data taken with FS007 and the superK laser at 495 nm. The prompt peak at approximately 250 ns in the time residual spectrum, shown in blue, therefore corresponds to the hits for which a straight line path including refraction best matches the path in reality. The offset of this peak from a time residual of zero simply comes from a trigger delay which is not taken into account in the time residual calculation. As previously discussed in Section 4.2, an extra delay is applied to the trigger issued by SMELLIE at TUBii to make sure that light is injected into the detector during the the EXTA trigger window. This delay is tuned so that light injection occurs approximately half way through the trigger window. Paths that involve scattering or reflection have greater time residuals than those in the prompt peak since the equivalent straight line paths are shorter and therefore the time of flight subtracted is less than for the real path taken.

The exception to this is involves the boolean in the `LightPathCalculator` class called `SetELLIEReflect`. As discussed in Chapter 2, this considers light paths that involve a reflection from the side of the AV closest to the fibre, henceforth called the nearside of the AV. Setting this boolean to `True` means that light reflected from the nearside of the AV is correctly modelled and therefore arrives promptly in the time residual spectrum. The difference this boolean makes to the time residual spectrum is summarised in Fig. 5.4.

The most obvious difference between the two spectra is that, when `SetELLIEReflect` is set to `True` (shown in blue), the shoulder to the right of the prompt peak disappears. This shoulder consists largely of AV reflected photons which are offset from
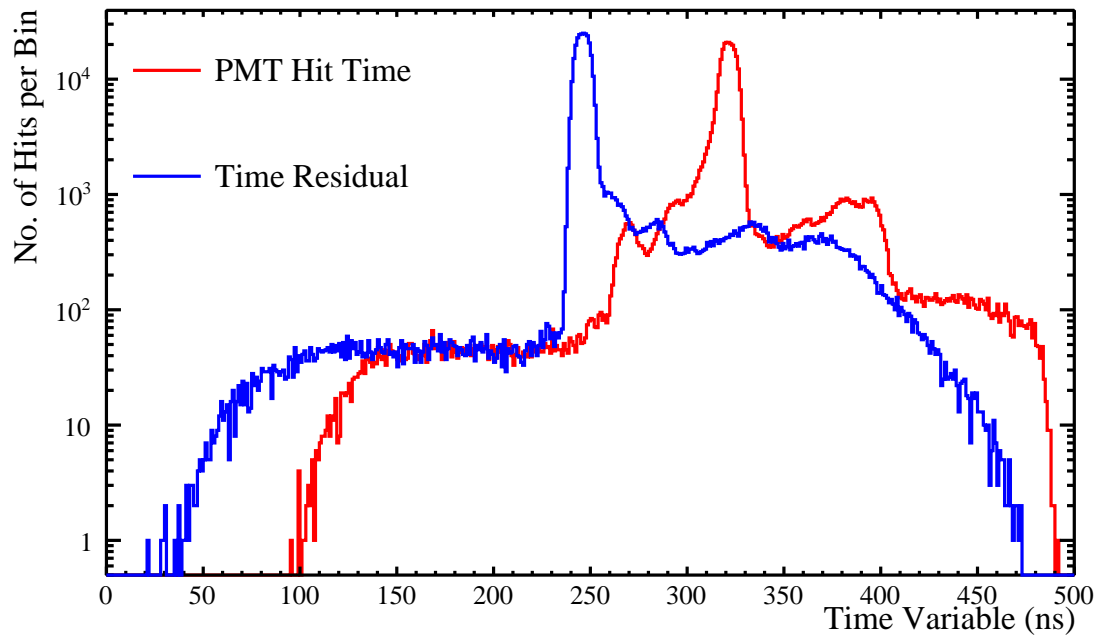
Figure 5.3: The calibrated hit time spectrum for the PMTs in the SNO+ detector summed over a subrun of SMELLIE data (495 nm, FS007) in red. Overlaid is the time residual spectrum for the same subrun of data in blue.
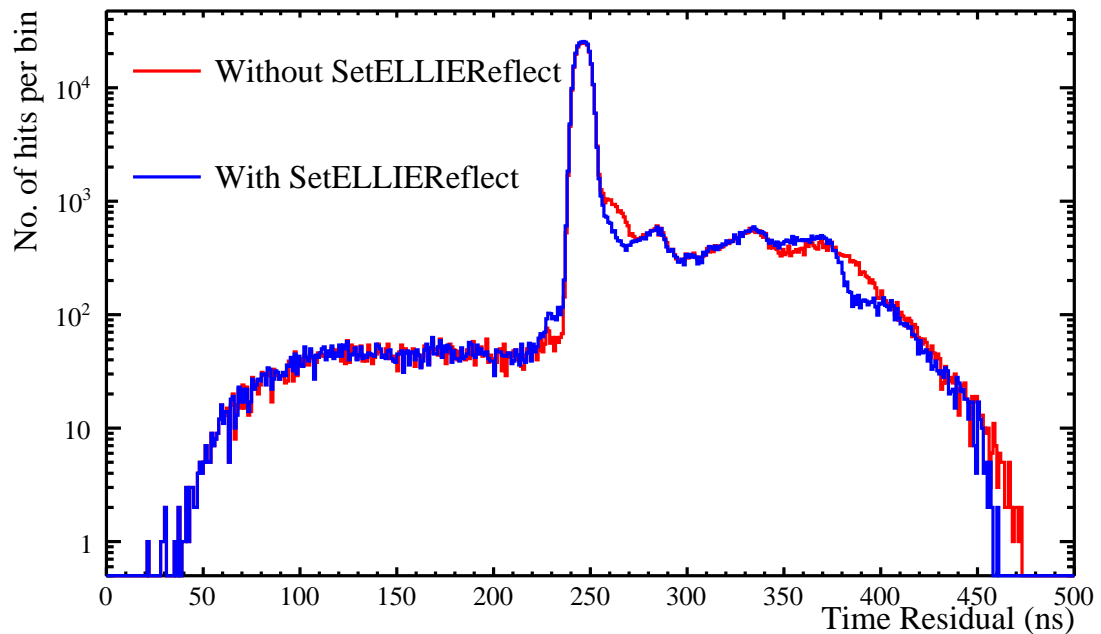
Figure 5.4: Time residuals calculated with (blue) and without (red) the `SetELLIEReflect` functionality used to account for reflections that occur from the nearside of the AV. These histograms are also created from 495 nm, FS007 data.

the prompt peak by a small time residual due to the difference between their true flight path (reflected from the near side of the AV to the fibre) and the flight path calculated by `LightPathCalculator` (straight line from fibre to PMT). This shoulder is absorbed into the prompt peak when such paths are properly modelled as the two paths are now the same and therefore there is no flight time difference between them.

The time residual spectrum is easiest to understand when the hits are broken down into categories of optical effects, such as in Fig. 5.5. In this plot, of a simulation of the same subrun as in the previous figures is shown. Recording of tracking information has been enabled in the simulation and using this information, the photons have been split into individual histograms according to the optical effects they underwent before entering a PMT. For example, the red histogram represents photons which have directly travelled to the PMT they entered after leaving the fibre position without undergoing optical effects other than refraction through the AV - no scattering or reflection. This red histogram has a clear prompt peak and smaller early and late light peaks, whereas the green histogram (photons which are reflected once by PMT glass before entering a PMT) has orders of magnitude fewer prompt photons, but dominates the late (post after pulsing) light spectrum. These categories of optical effects will be explored further in Section 7.1.1.

When comparing SMELLIE time residuals in data to simulation, a discrepancy is apparent in the time residual spectrum, shown in Fig 5.6 for the superK laser at 495 nm and in Fig 5.7 for the PQ375 laser.

Clearly, for the PQ lasers, the trigger has been delayed by different amount than the superK laser and the simulation does not take this into account, as evidenced by the approximately 100 ns difference in the prompt peak of the time residual spectra in Fig. 5.7. Even for the superK, the delay is not completely accurate. However, for the analysis, this difference in delay is not problematic: provided the delay is identical for every event, it can simply be subtracted from every hit, centring the prompt peaks around zero.

However, from the Gaussian fits which have been applied to all four of the prompt peaks, it can be seen that the widths of the prompt peaks are consistently different in data and simulation. This is a concern as this cannot be corrected for in analysis unlike an offset. For the superK laser, the Gaussian sigma in data is 3.502 ± 0.004 ns, compared to 1.874 ± 0.003 ns in simulation. (These values are from subruns of FS007 at 495 nm.) For the PQ375 (using FS007 again), the Gaussian sigma in data is 3.742 ± 0.004 ns and 1.766 ± 0.002 ns in simulation. Therefore, for both the superK
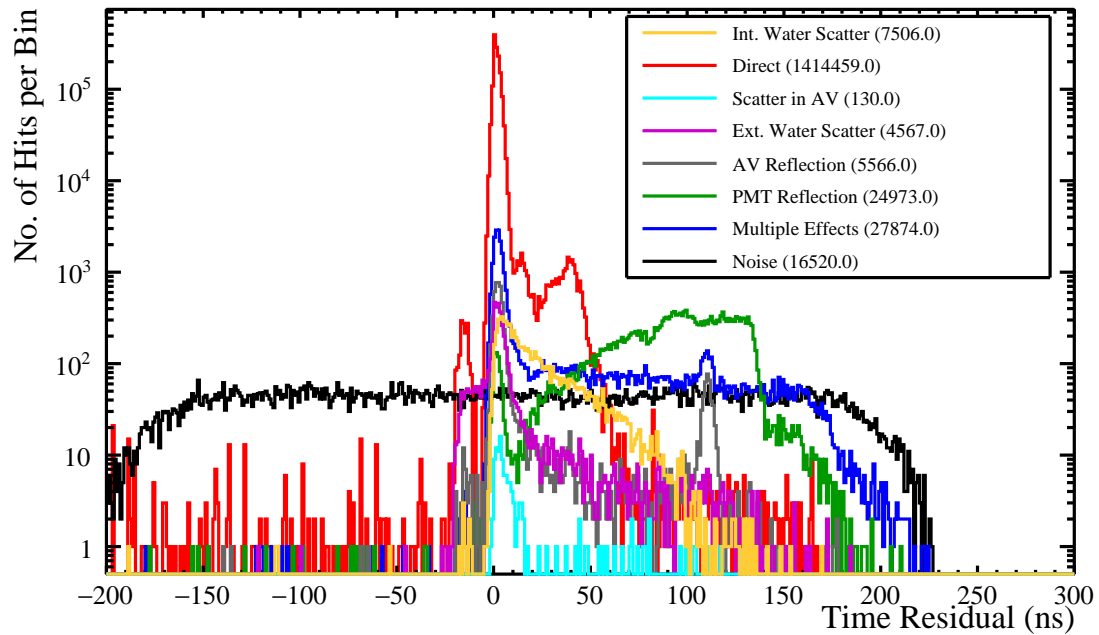
Figure 5.5: The time residual spectrum for FS007 at 495 nm in simulation. Tracking information has been enabled. Using this, the photons have been categorised into different categories according to the optical processes they have undergone before reaching PMTs. The numbers in the legend refer to the number of photons in each optical category.
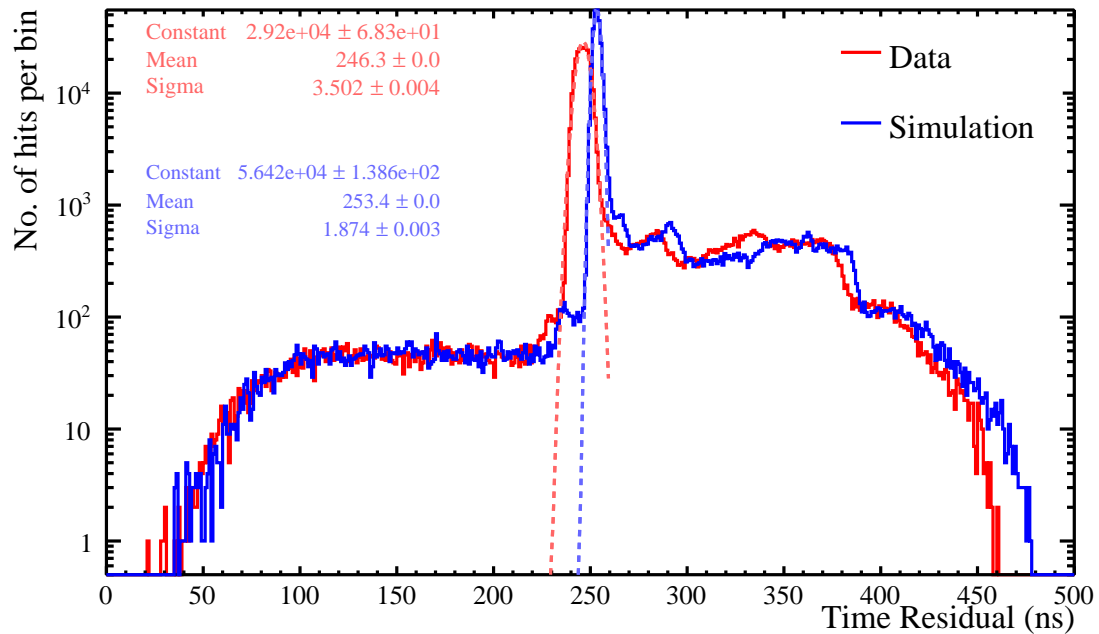
Figure 5.6: Comparison of time residuals calculated for a data subrun (red) and a matching simulation (blue) of the superK laser at 495 nm. Both the prompt peaks have been fitted with Gaussians and the results from those fits are shown in the top left of the figure.
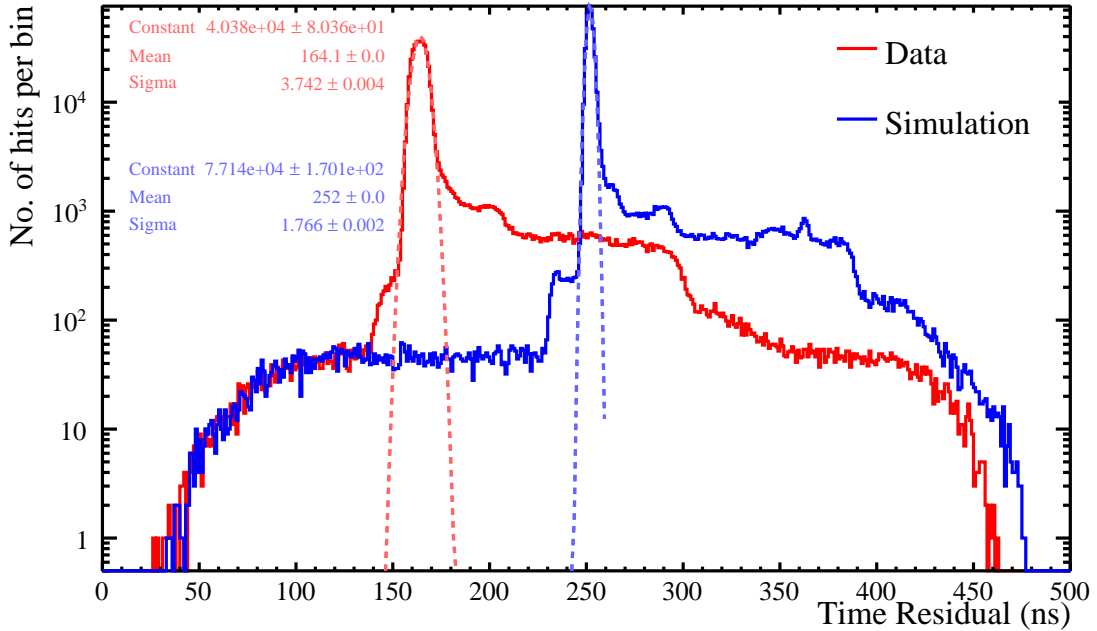
Figure 5.7: Comparison of time residuals calculated for a data subrun (red) and a matching simulation (blue) of the PQ375 laser. Both the prompt peaks have been fitted with Gaussians and the results from those fits are shown in the top left of the figure.

and the PQ lasers, there is approximately a 2 ns width discrepancy in the time distribution between data and simulation.

One possibility for the cause of this discrepancy is a trigger jitter in the trigger system for the lasers. However, the superK and PQ lasers have different trigger paths at the SMELLIE rack, as discussed in Chapter 4, making a consistent discrepancy in both systems extremely unlikely.

SMELLIE is the only calibration system which uses master mode as well as using TUBii to delay the trigger signal before it is sent onto the MTC/D to issue a detector trigger. (The trigger path is discussed more Chapter 4.) Therefore, it was a possibility that this process applies a constant jitter to the trigger signal. This would explain why only the SMELLIE analysis encounters this discrepancy.

This possibility was investigated by using hit time data to set the zero time (t_0) for each event rather than the trigger time. This was done by choosing a number of PMTs closest to the beamspot centre position on the PSUP. Then the second hit in this group of PMTs is chosen as the new t_0, as this seems less likely to be a noise hit than the first.

The effect of the number of PMTs selected and the hit used as the t_0 are examined in Figs. 5.8 and 5.9 respectively. The Gaussian sigmas are not affected by either choice
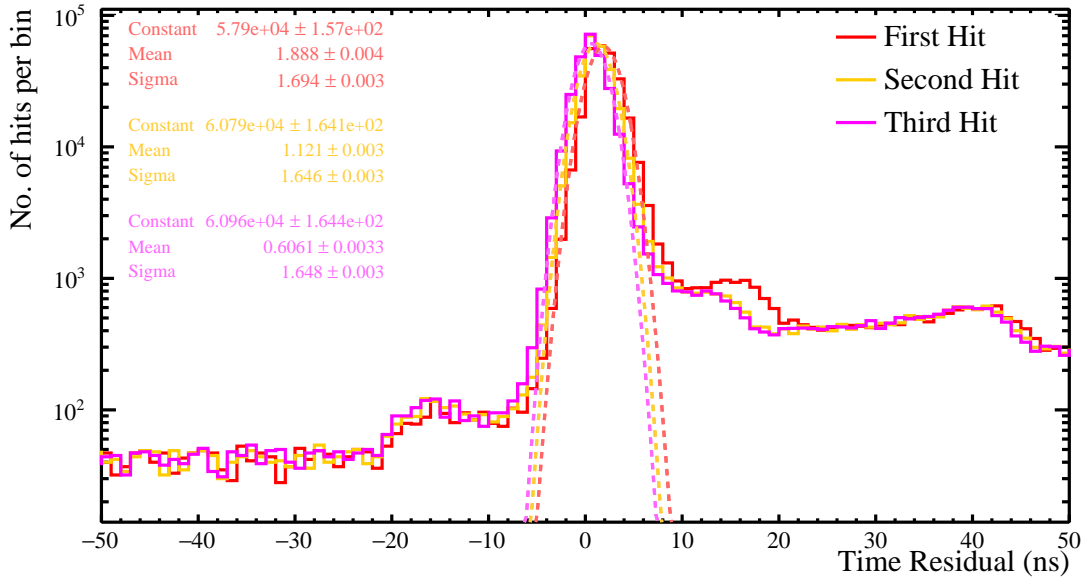
Figure 5.8: Comparison of time residuals using a data driven t_0 calculated for a data subrun of the superK laser at 495 nm. The 10 PMTs closest to the beamspot centre have been selected and the first (red), second (orange) and third (magenta) hits in this group of PMTs have been used as the t_0 for each event. The prompt peaks have been fitted with Gaussians and the results from those fits are shown in the top left of the figure.

(within 0.05 ns). The mean is shifted by the choice of which hit is used; however, as previously discussed, this can be corrected for in the analysis. For future use, the number of PMTs has been chosen to be 20 and, as already mentioned, the second hit has been used.

Figs. 5.10 and 5.11 compare this new time residual using the data driven t_0 to the simulations. This demonstrates that the increased width of the prompt peak is due to jitter in the trigger system, as using a data driven t_0 rather than the EXTA trigger time as a t_0 removes the excess width of the prompt time residual peak. Suspicion specifically falls on the delay applied to the trigger signal by TUBii as this is the part of the trigger path shared by all the lasers and not by any other calibration system. Enquires were made to the TUBii team and it was discovered that the delay is synchronised to a 100 MHz clock [107]. This would be expected to introduce a flat 10 ns jitter to the EXTA trigger as each trigger pulse is synchronised to the next clock tick.

It is important to realise that applying a flat 10 ns jitter to the trigger pulses would not result in a Gaussian distribution with a sigma on the order of 10 ns for the prompt peak. Rather the effect of convoluting the delay with the true prompt peak must be
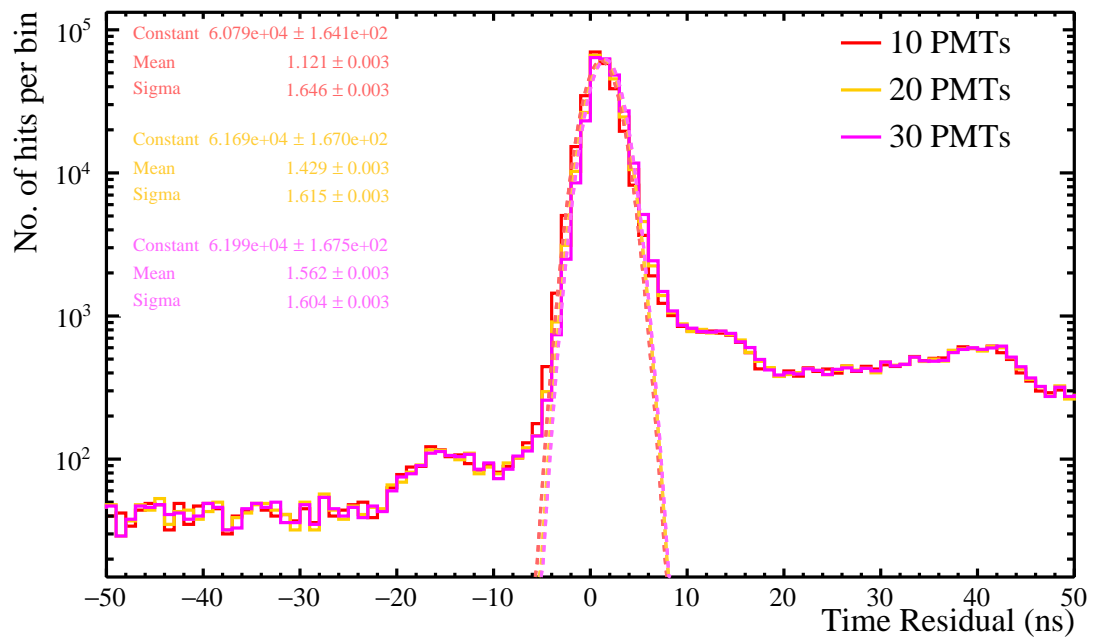
Figure 5.9: Comparison of time residuals using a data driven t_0 calculated for a data subrun of the superK laser at 495 nm. The 10 (red), 20 (orange) and 30 (magenta) PMTs closest to the beamspot centre have been selected and the second hit in this group of PMTs has been used as the t_0 for each event. The prompt peaks have been fitted with Gaussians and the results from those fits are shown in the top left of the figure.
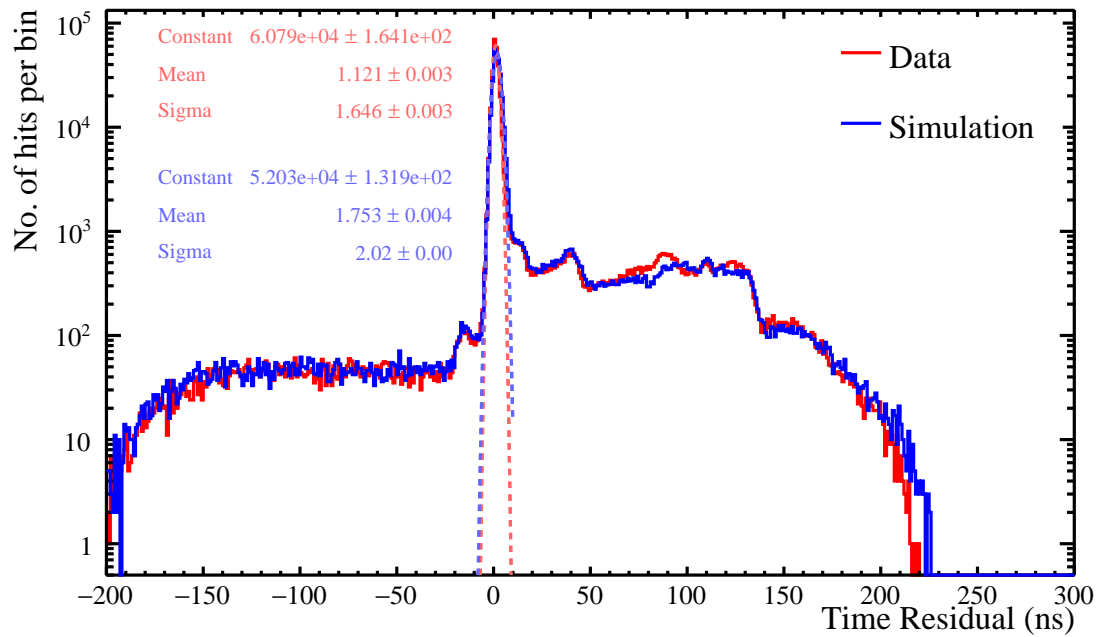
Figure 5.10: Comparison of time residuals using a data driven t_0 calculated for a data subrun (red) and a matching simulation (blue) of the superK laser at 495 nm. Both prompt peaks have been fitted with Gaussians and the results from those fits are shown in the top left of the figure.
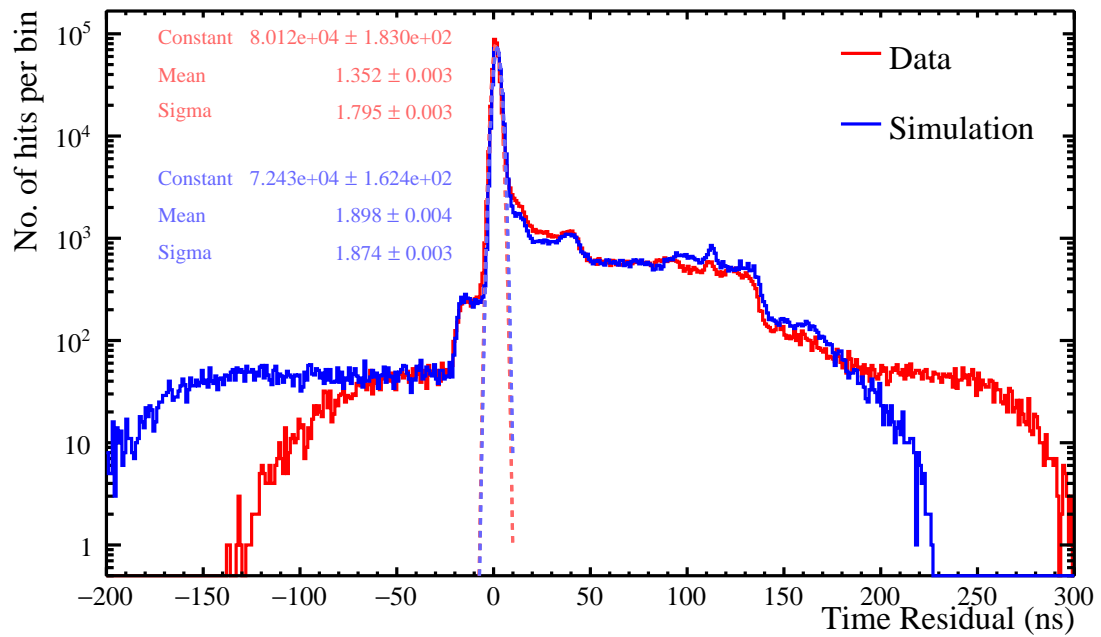
Figure 5.11: Comparison of time residuals using a data driven t_0 calculated for a data subrun (red) and a matching simulation (blue) of the PQ375 laser. Both prompt peaks have been fitted with Gaussians and the results from those fits are shown in the top left of the figure.
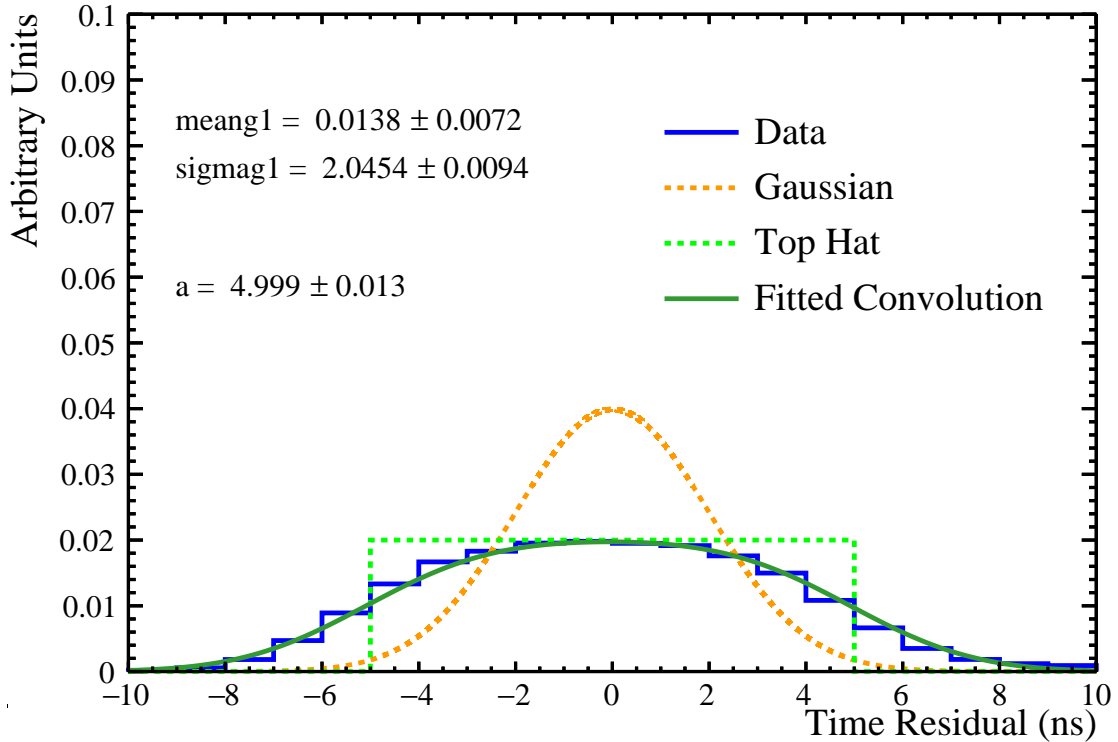
Figure 5.12: Convolution of a Gaussian distribution (yellow, dashed) and a top hat function (green, dashed) designed to model the jitter added to the EXTA trigger by the TUBii delay being synchronised to a 100 MHz clock. The resultant convolution (green, solid) has been fitted to the prompt time residual peak in data (blue, solid). The fit results for the original variables of the top hat function (half-width - a) and Gaussian (mean - meang1 and sigma - sigmag1) once the convolution has been fitted are shown in the top left of the figure.

considered to understand whether synchronisation to a 100 MHz clock is plausible as the source of the jitter in the EXTA trigger.

To investigate this, a Gaussian distribution (to model the true time residual spread in the prompt peak from the TTS of the PMTs and the timing distribution of the lasers) has been convolved with a top hat function (to model the jitter from the TUBii delay). The convolution has then been fitted to the prompt time residual peak in data. The result is shown in Fig. 5.12. The fit results in terms of the variables of the original top hat and Gaussian distributions are shown in the top left (meang1 and saigmag1 being the mean and sigma of the Gaussian and a the half width of the top hat).

It can be seen that fitting this convolution to the prompt time residual peak of the data results in a Gaussian distribution with a sigma of 2 ns and a top hat function with a half-width of 5 ns. This supports the hypothesis that the width of the prompt

peak in data is, indeed, caused by the trigger delay being synchronised to a 100 MHz clock.

Therefore, not only can this trigger jitter be corrected for in water data through creating a data driven t_0, but the source of the trigger jitter is understood. This means that while the detector is being filled with scintillator, a hardware fix can be implemented so that using a data driven t_0 will not be necessary during future SMELLIE analyses. This hardware fix will take the form of implementing a truly asynchronous external trigger delay in TUBii. For the analysis presented in this work, however, it is too late to make hardware changes to TUBii and therefore, a data driven t_0 is used, in the form of the second hit in the 20 PMTs closest to the beamspot centre.

5.4 Rings

When the commissioning of SMELLIE began, it was immediately apparent that there are unexpected features present in the spatial distribution of detector NHits. For certain fibres, there is a clear ring-like structure around the central beamspot, as shown in the detector flat map (the PSUP unfolded into 2D space) in Fig. 5.13 for FS107. For comparison, Fig. 5.14 shows data from a subrun of FS007, in which no ring structure can be seen. These rings were initially identified by creating such flat maps and checking the NHits distributions by eye. This identified rings present in data taken using FS107, FS155, FS125 and FS093. The injection points for these four fibres are distributed throughout the detector and include fibres for at least one of each pointing angle.

The rings are located at an α angle of approximately 0.3 rads. This can be seen by creating histograms of the p.e. per PMT as a function of α. An example histogram is shown in Fig. 5.15, which compares data from a fibre in which a ring is visible (FS107) to data from a fibre without a visible ring (FS007). The two subruns come from the same run (106846). The histogram for each fibre has not only been normalised by the number of PMTs in each bin, but, in addition, the number of p.e. in each PMT has been divided by the number of p.e. in the same PMT in a simulation that models SMELLIE as an isotropic light source. This normalisation method accounts for effects caused by the detector geometry, such as PMT angular acceptances. See Chapter 6 for a full discussion of this normalisation technique. The ring location of α equal to approximately 0.3 rads is consistent across the four fibres.
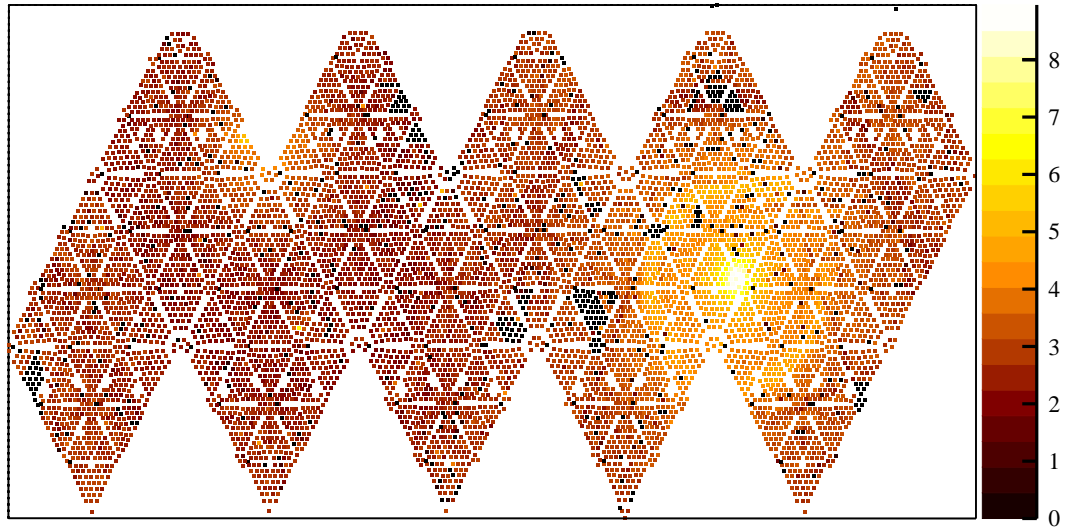
Figure 5.13: A flat map of the detector for a SMELLIE subrun; every point is a PMT. The colour represents the natural logarithm of the NHits per PMT, summed over the subrun. A ring is visible around the beamspot on the right hand side of the plot. (The bright area in the top left is the reflection of the beam from the side of the AV closest to the fibre.) This data was taken using FS107, at a wavelength of 415 nm.
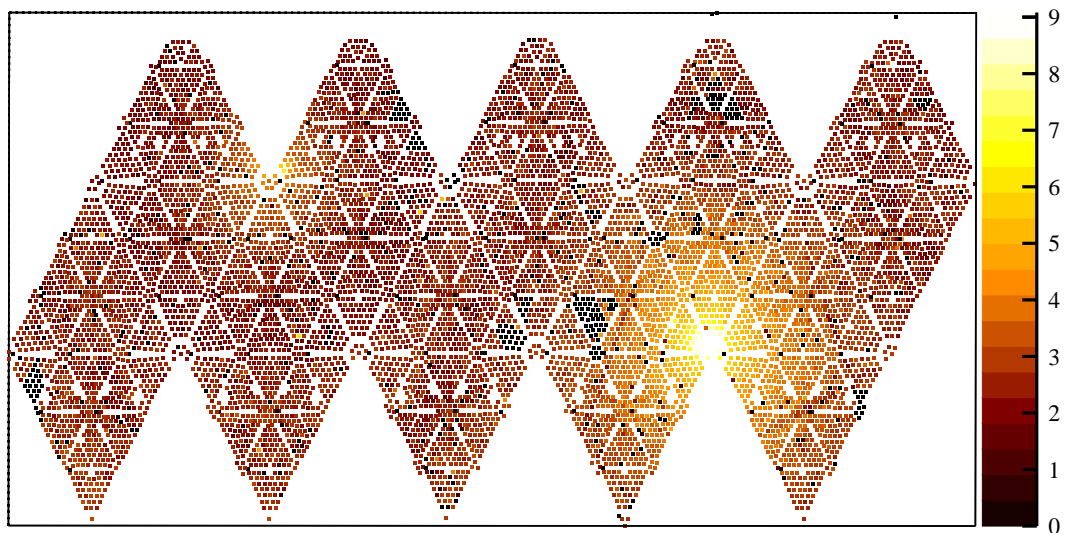
Figure 5.14: A flat map of the detector for a SMELLIE subrun; every point is a PMT. The colour represents the natural logarithm of the NHits per PMT, summed over the subrun. No ring is visible. This data was taken using FS007, at a wavelength of 415 nm.
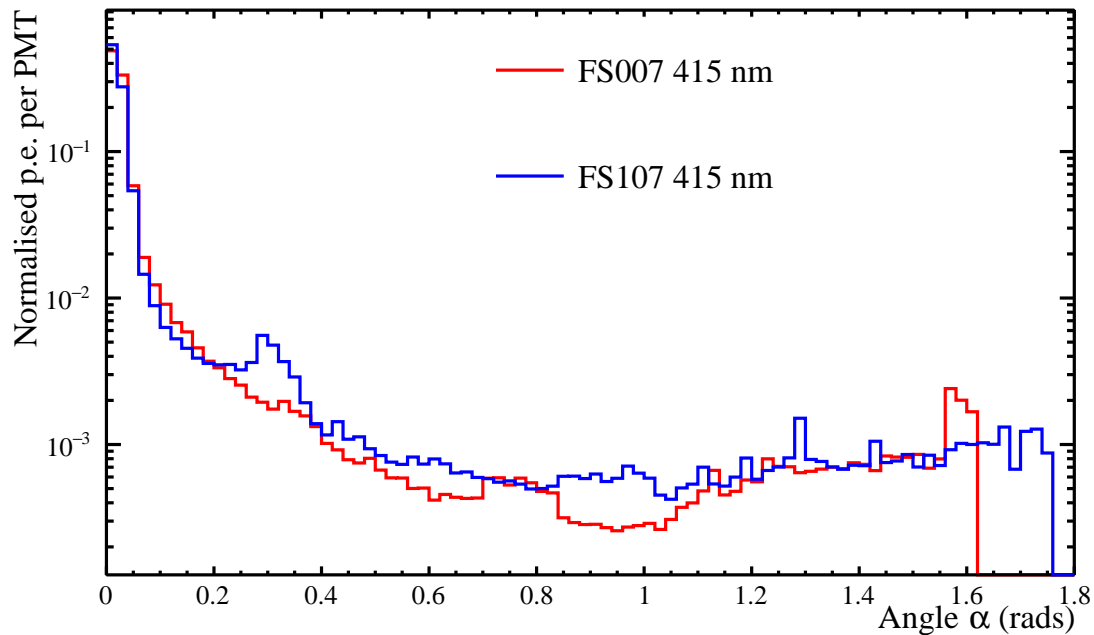
Figure 5.15: The p.e. per PMT in the detector during two separate SMELLIE subruns in run 106846. Both subruns have been normalised by isotropic simulations and both used the superK laser with a central wavelength of 415 nm. The data from FS107 (blue) shows a peak at approximately 0.3 rads, whereas the data for FS007 (red) does not.

Using histograms such as Fig. 5.15, some of the 11 other fibres seem to have features at approximately 0.3 rads that could indicate rings, notably shoulders on the falling edge of the beamspot. This is true for FS137 and FS237. Additionally, during this data taking period, FS193 and FS293 were not available (the secondary fibre switch had not been successfully installed yet). In later data, it became apparent that the beam of FS293 also has a ring feature.

There are several possibilities for the cause of these rings. The potential cases can be split into two categories - either the ring is part of the photon emission profile of the fibre or the ring is caused by an effect in the detector, after the light is emitted from the fibre. Determining which category the rings belong to is crucial as this determines whether the rings need to be included in the modelled angular emission profile of the SMELLIE fibres (discussed in Chapter 6) or whether there is some error in the current detector simulation, as no trace of a ring has ever been seen in a SMELLIE simulation.

Firstly, causes linked to a detector effect are discussed. Given that the ringed beams are emitted from fibres which are located throughout the detector, the possibility of the rings being caused by anything specific to the beam location, such as a defect on

the surface of the AV, is highly unlikely.

The possibility of the rings being some form of reflection was discounted after comparing the time residuals of the hits which form the rings to those of the central beamspot. It was found that the peak time residual for PMT hits in the ring was the same as for the beamspot, meaning that the photons causing the hits in the ring must have travelled from the fibre to the PMTs, without undergoing additional delays.

Another possible cause for the rings is the presence of air bubbles trapped in the aperture of the collimators (discussed in Chapter 4). This is unlikely for several reasons. Firstly, again, the ringed beams originate from fibres distributed across the detector in z-coordinates, meaning that bubbles producing identical ring patterns would have to form for fibres both pointing upwards and downwards. Considering that the cavity is filled gradually from the bottom, this does not seem plausible. In addition, there is a limited amount of SMELLIE data that was taken prior to a completely water-filled detector. One of the fibres which was fired through air was FS155. This data is extremely noisy and not all of the crates were at HV. However, there does appear to be some ring-like structure present.

Perhaps the most obvious possibility is that the rings are diffraction rings, likely from an aperture in the collimator. However, the behaviour of the rings does not match expected diffraction behaviour. To first order, the aperture required to cause an Airy ring with $\alpha \approx 18°$ at 495 nm, using $d = \frac{1.635\lambda}{\sin\alpha}$, where d is the diameter of the aperture, would be 3 μm. There is no obvious aperture in the detector, fibre or collimator with a diameter on that order. For instance, the aperture of the core of the fibre is 50 μm.

Diffraction has well understood wavelength dependant behaviour. This means that the behaviour of the ring as a function of wavelength can be compared to the expectations for diffraction as a further test. If the ring is caused by diffraction and an aperture of 3 μm (as calculated above), the change in α at the extremes of the wavelength range is calculated. For 415 nm, α would be approximately 13°, compared to 595 nm and an α of 19°. This is a change of $\sim 6°$. As an estimation, for a 0° fibre, a PMT at the beamspot centre is approximately 1° in diameter when viewed from the fibre position. Therefore, a change in the ring position by 6° (or approximately 0.1 rad) should be visible in the detector.

The position of the ring is then examined in data as a function of wavelength. An example of this is shown in Fig. 5.16 for FS107 at 415 and 595 nm. As clearly shown in this plot, the position of the ring is consistent at both wavelengths and therefore
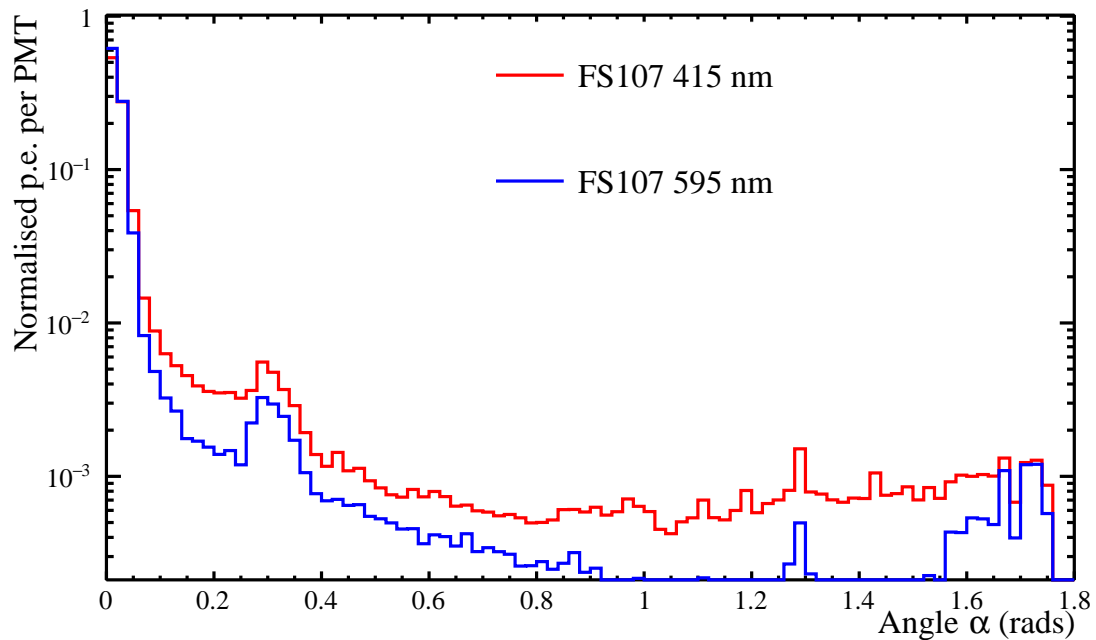
Figure 5.16: The p.e. per PMT in the detector during two separate SMELLIE subruns, one using the superK at 415 nm (red) and the other at 595 nm (blue). Both subruns have been normalised by isotropic simulations and both used FS107 as the beam emitted from this fibre shows a ring feature.

does not exhibit the expected behaviour of a diffraction ring. As before, this plot uses normalisation by isotropic simulations.

As a check, if it is assumed that the ring is caused by diffraction and that it is not possible to see the ring move more than a PMT over this wavelength range, then it is possible to calculate what α would have to be at 495 nm for this to be true. First, a value of d must be calculated for which any movement of the ring centre between 415 nm and 595 nm cannot be discerned: at $d = 25$ μm, the change in α between 415 and 595 nm is smaller than a degree. If $d = 25$ μm, α at 495 nm would be $\sim 2°$. This is not where the ring is observed. This provides further evidence that these ring features do not behaviour as diffraction rings.

The intensity of light within the ring as a function of wavelength was also examined to further understand the behaviour of the ring structures. In order to do this, histograms normalised by an isotropic simulation as shown in Fig. 5.16 are created, at 10 steps in wavelength between 415 and 595 nm. A ratio is then created of the normalised p.e. in the α region of the ring (defined as $0.2 - 0.4$ rads) over the normalised p.e. in the beamspot ($0 - 0.2$ rads). This ratio is linearly fitted as a function of wavelength for each fibre and the gradients and intercepts of these fits are shown in Figs. 5.17
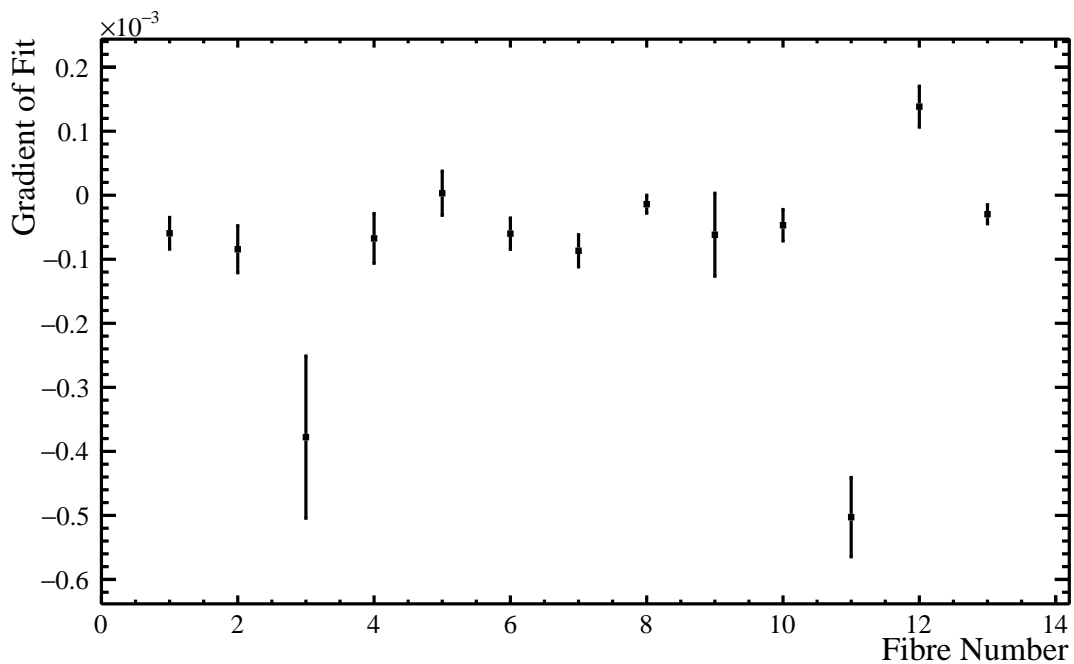
Figure 5.17: The gradient of the fit of the ratio of p.e. in the ring region to p.e. in the beamspot as a function of wavelength for each fibre. The error bars are the errors of the fit, arising from Poisson errors in the number of p.e. in each region at each wavelength for each fibre. The fibre numbers follow this order: FS007, FS107, FS207, FS025, FS125, FS225, FS037, FS137, FS237, FS055, FS155, FS255, FS093.

and 5.18. FS193 and FS293, as discussed earlier in this section, were not part of the commissioning data used for this investigation.

These plots show that, while there is clear variation in the gradients and intercepts of these fits, the subruns in which rings are present (in particular, fibre numbers 2, 5, 11 and 13) show no consistent behaviour that is different from the subruns which do not feature rings. Therefore, the conclusion is that the behaviour of the intensity of the ring relative to the beamspot is no different to the same region when absent of a ring.

All of these investigations indicate that the ring features are part of the emission profile of light from the fibre. Exactly what is causing them remains unknown. However, if they are part of the emission profile, they undergo no additional optical processes to the direct beamspot and therefore can simply be included in the modelled angular emission profile. In order to do this, detector data must be used to create the angular emission profiles, as the profiles measured ex-situ did not extend far enough in α to include the ring region. The creation of these profiles and accurate generation of a SMELLIE beam is the subject of Chapter 6.
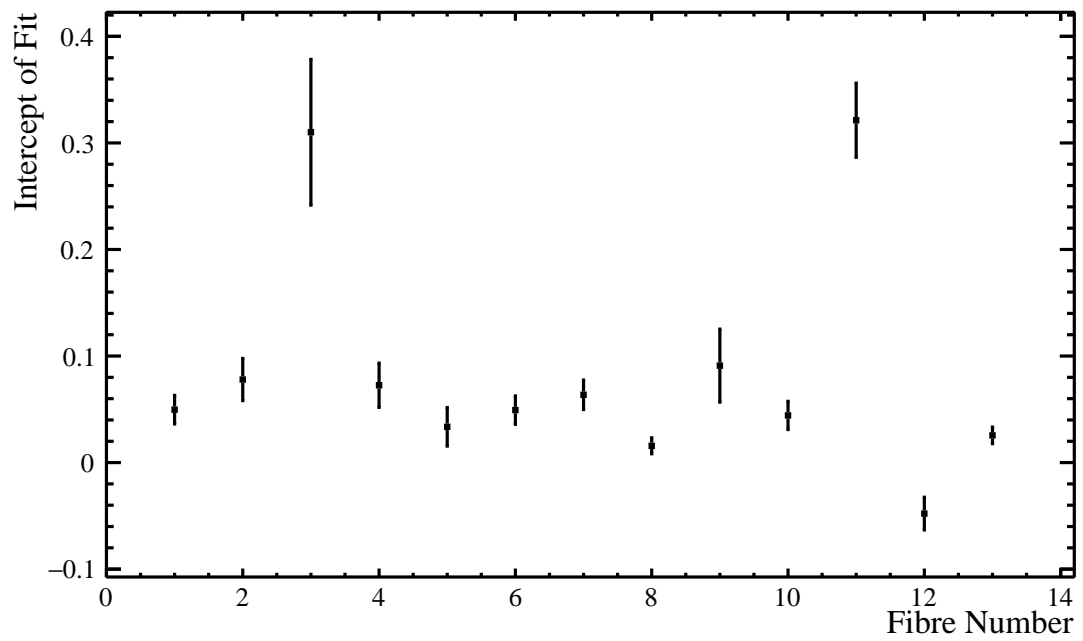
Figure 5.18: The intercept of the fit of the ratio of p.e. in the ring region to p.e. in the beamspot as a function of wavelength for each fibre. The error bars are the errors of the fit, arising from Poisson errors in the number of p.e. in each region at each wavelength for each fibre. The fibre numbers follow this order: FS007, FS107, FS207, FS025, FS125, FS225, FS037, FS137, FS237, FS055, FS155, FS255, FS093.
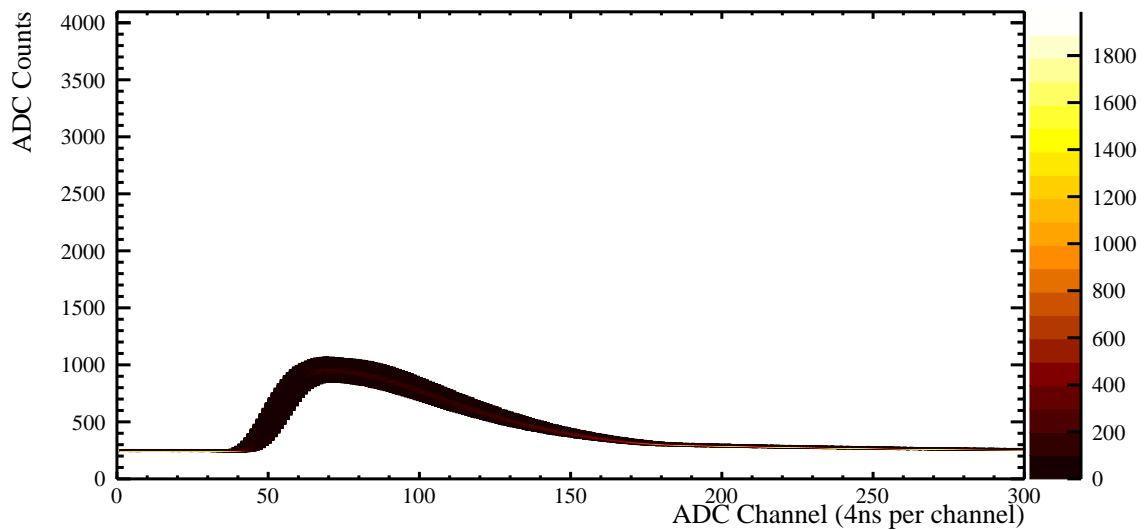
Figure 5.19: Each CAEN trace from the MPU for a subrun of SMELLIE data using the superK at 495 nm overlain on each other. The colour represents the number of traces overlain at each point.

5.5 CAEN Traces

The SMELLIE MPU and its output in the form of CAEN traces (see Section 4.1.4) give an invaluable shot by shot measurement of the energy of each laser pulse and hence how many photons are contained in it. In this work, the CAEN information is mainly used as a quality assurance check and to evaluate systematics. Therefore, it is not fully exploited as it is theoretically possible to use the CAEN traces to implement a shot by shot p.e. count correction (as discussed in [106]) as an extension to the already existing multi-hit correction.

As a quality assurance check, the CAEN traces for each subrun are recorded, allowing the analyser to check these for unusual behaviour. A typical set of CAEN traces is shown in Fig. 5.19. This particular example is a subrun of the superK laser firing at 495 nm.

However, as is shown in Figs. 5.20 and 5.21, the CAEN traces can be significantly less stable than in Fig. 5.19. This is typically caused by the laser operating near its lasing threshold; there are large variations in the number of photons emitted per shot. These two examples both used PQ lasers: Fig. 5.20 shows a subrun in which the PQ405 was used and Fig. 5.21 the PQ440.

The CAEN has a range of 0 - 4095 ADC counts. Once the trace reaches 4095 counts (called a saturated CAEN trace), most information from the MPU concerning the shot by shot photon number is lost. Another example of this is shown in Fig. 5.22
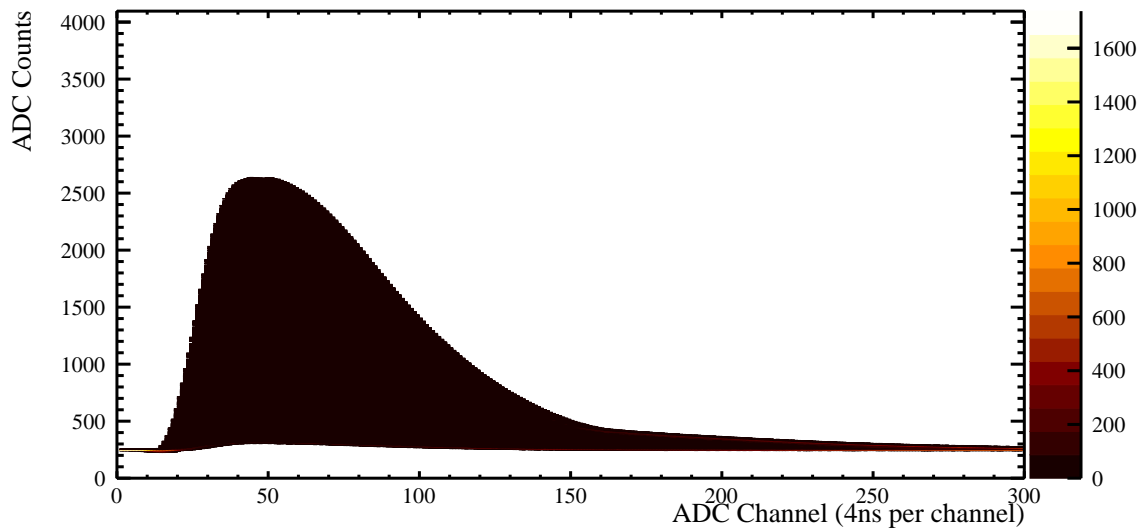
Figure 5.20: Each CAEN trace from the MPU for a subrun of SMELLIE data using PQ405 overlain on each other. The colour represents the number of traces overlain at each point. This is an example of laser instability.
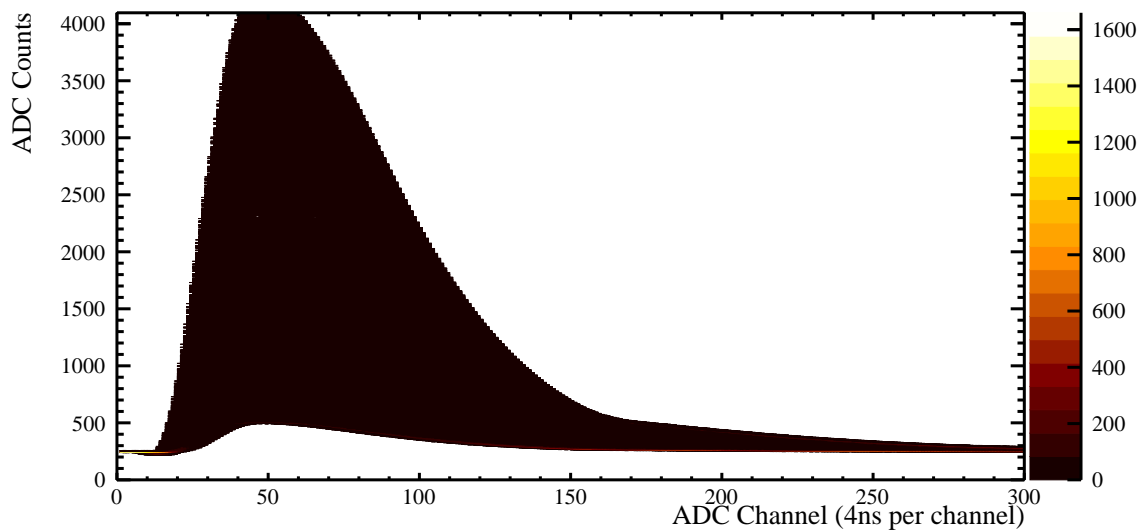
Figure 5.21: Each CAEN trace from the MPU for a subrun of SMELLIE data using PQ440 overlain on each other. The colour represents the number of traces overlain at each point. This is an example of laser instability.
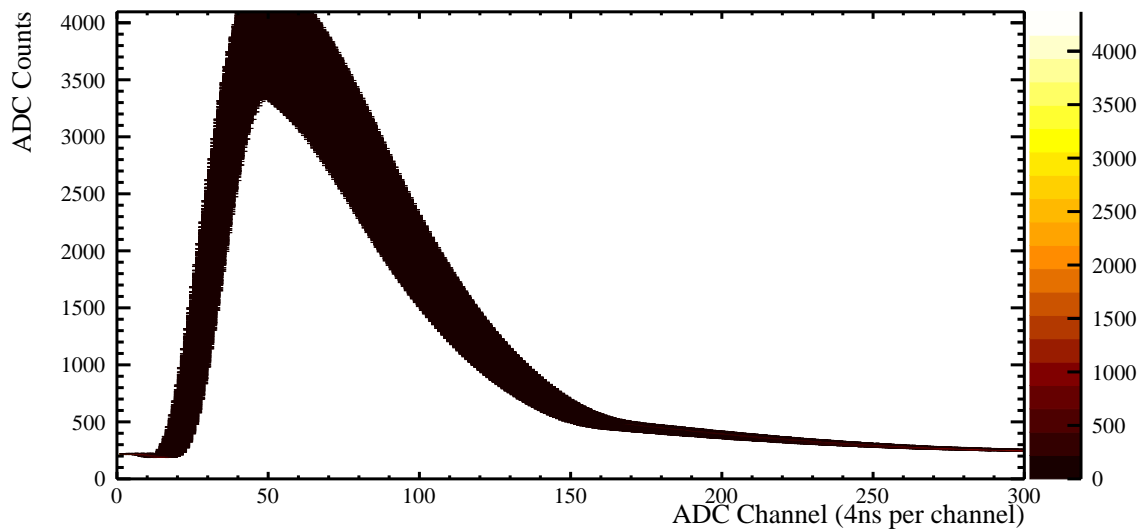
Figure 5.22: Each CAEN trace from the MPU for a subrun of SMELLIE data using PQ405 overlain on each other. The colour represents the number of traces overlain at each point. This is an example of saturated CAEN traces.

in which the PQ405 laser was used. In such a case, decreasing the gain of the MPU for a future subrun that is otherwise identical would be sufficient to make the CAEN traces useable.

However, in other subruns, such as in Fig. 5.21, the laser instability is so great that some traces are saturated while others are almost flat. This means that the gain cannot be adjusted to compensate. The only option is to increase the attenuation of the attenuator following the laserhead (see Section 4.1.1) and operate it at a higher intensity setting to produce the same shot energy. However, this is not always possible as a range of shot energies are required for optimal operating of SMELLIE and if the attenuation is set too high, then achieving the highest shot energy setting (discussed in Section 5.6) will no longer be possible.

The attenuation adjustment as described in the previous paragraph was done for two of the PQ lasers: PQ405 and PQ440. This was done in two stages, meaning that the attenuations were not constant throughout the commissioning of SMELLIE. This is one of the main reasons that run 114018 (and the runs taken during the same shift) are used both for the beam profiling (in Chapter 6) and the scattering analysis (in Chapter 7), as well as for some of the commissioning: these runs were taken after the final attenuation tuning and, therefore, when the lasers were as stable as possible.

The attenuators were not adjusted for PQ375 and PQ495. This is largely because there were operational problems encountered when attempting to adjust them and,

Figure 5.23: The height of each CAEN trace from the MPU for a subrun of SMELLIE data (using the superK at 495 nm) plotted as a function of event GTID.

due to the time restrictions on shifts underground, there was not sufficient time to satisfactorily resolve these issues.

These laser instabilities need to be probed. Therefore, the peak CAEN trace height minus the baseline voltage (henceform simply known as the CAEN height) is plotted as a function of the event GTID (global trigger ID), which acts as a proxy for event time. This is shown for various different subruns of SMELLIE data in Figs. 5.23 - 5.27. Fig. 5.23 and Fig. 5.24 show the CAEN heights for the same subrun as in Fig. 5.19. This is a subrun from run 114018, taken using the superK laser at 495 nm and is an example of stable laser behaviour.

However, Figs. 5.25 - 5.27 all show SMELLIE subruns in which the CAEN heights have a time dependence. These three examples are all subruns in which one of the PQ lasers was used (specifically, PQ440 and PQ495). It is notable that the PQ lasers suffer from instability more than the superK laser. This is one of the reasons that the superK laser is favoured in the scattering analysis and only PQ375 is used; the superK spectrum does not extend to this short a wavelength.

In order to investigate these behaviour patterns, plots of the integral under the CAEN trace and the baseline of the trace are shown in Fig. 5.28 and Fig. 5.29 for the same subrun as shown in Fig. 5.27. This shows that the integral of the CAEN traces shows the same behaviour as the height, as expected, and that the baseline does not
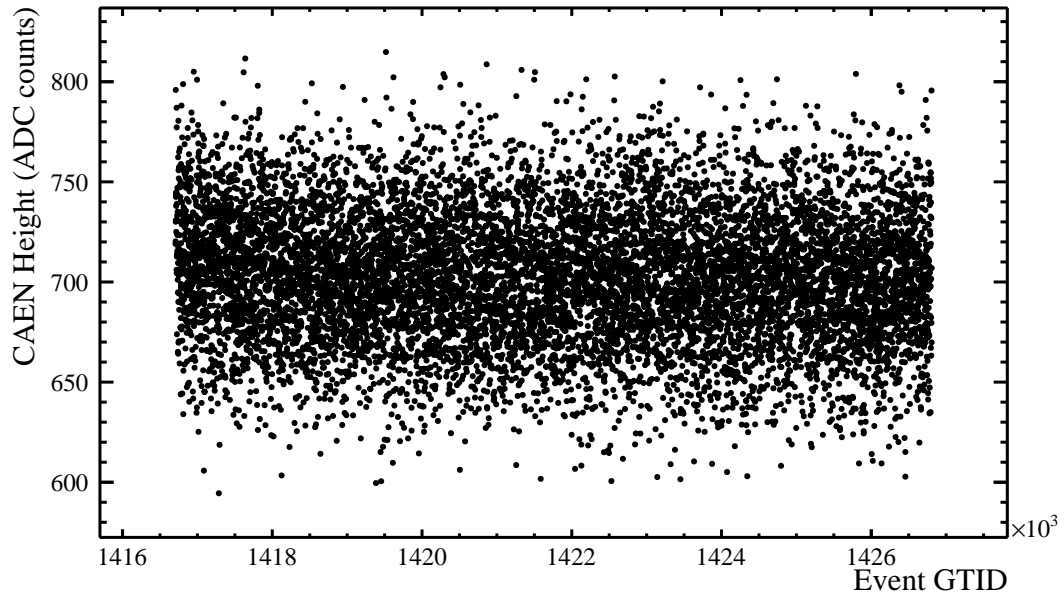
Figure 5.24: The height of each CAEN trace from the MPU for a subrun of SMELLIE data (using the superK at 495 nm) plotted as a function of event GTID. This is the same subrun as in Fig. 5.23, but zoomed in so the distribution of CAEN heights is visible.
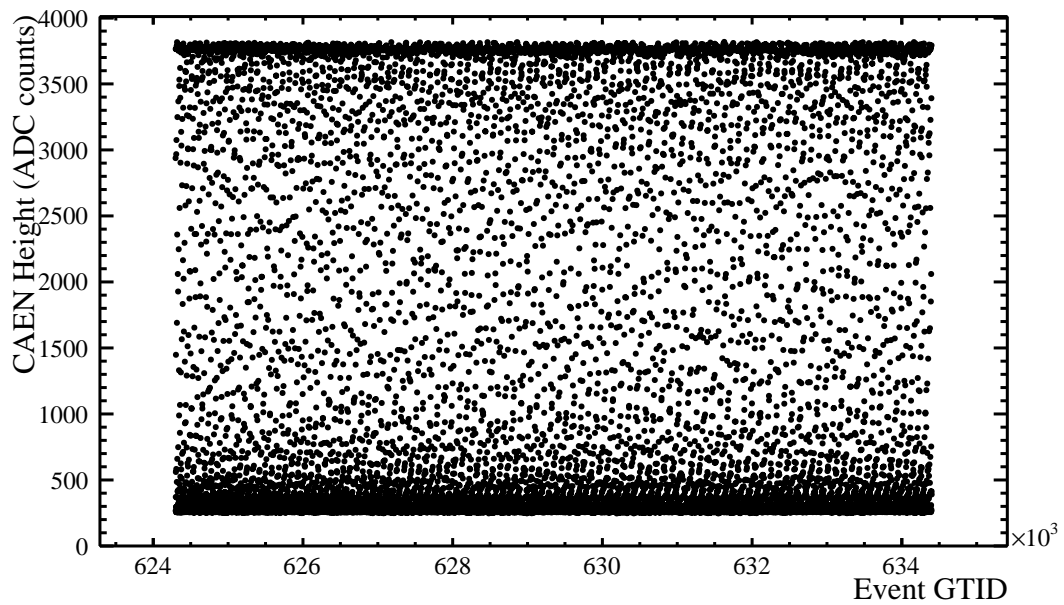
Figure 5.25: The height of each CAEN trace from the MPU for a subrun of SMELLIE data (using PQ440) plotted as a function of event GTID.
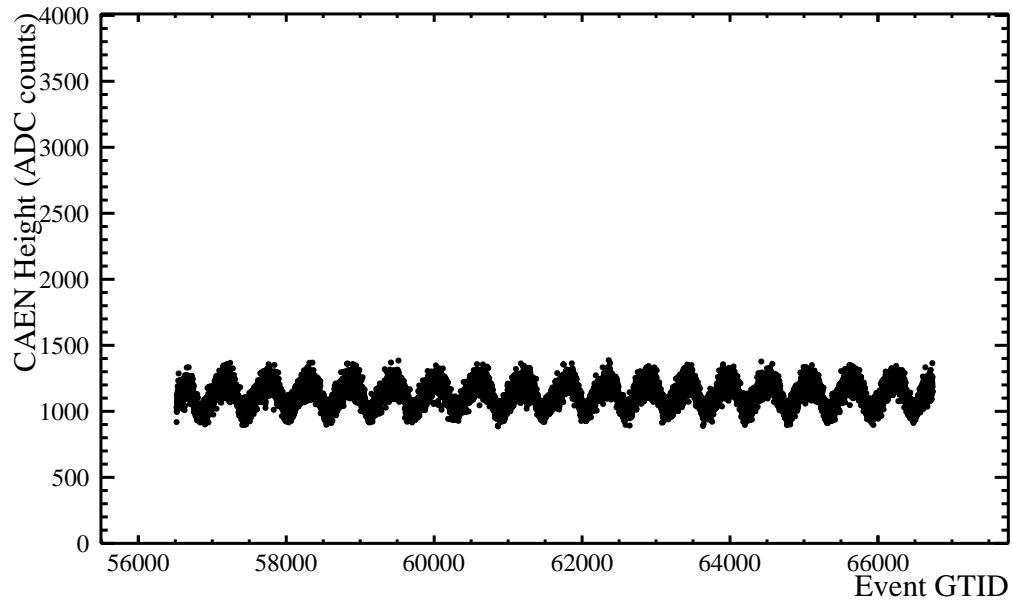
Figure 5.26: The height of each CAEN trace from the MPU for a subrun of SMELLIE data (using PQ495) plotted as a function of event GTID.
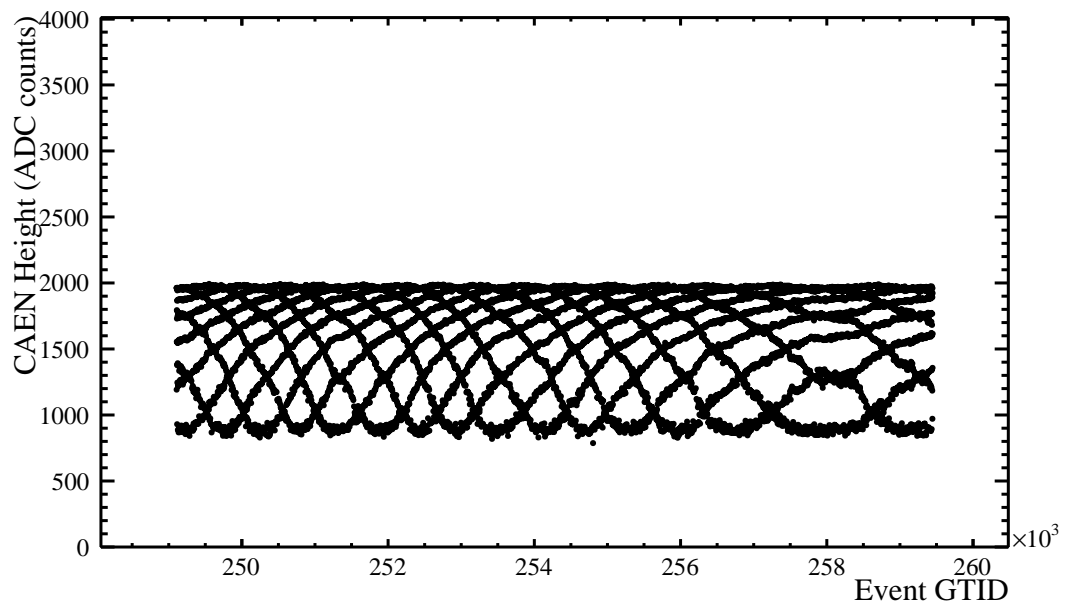
Figure 5.27: The height of each CAEN trace from the MPU for a subrun of SMELLIE data (using PQ440, though a different subrun to that in Fig. 5.25) plotted as a function of event GTID.
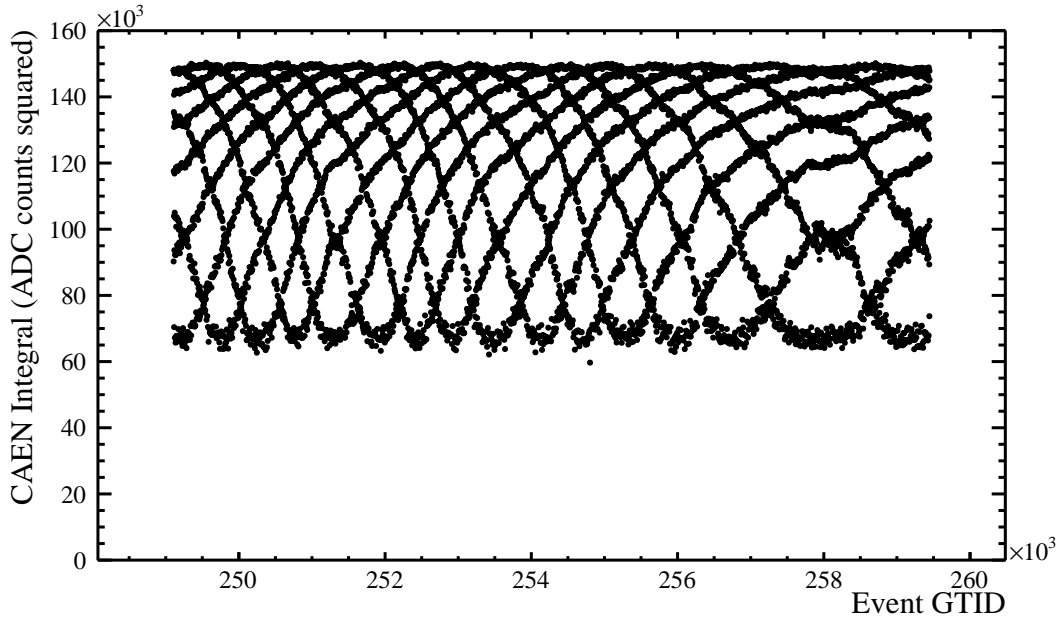
Figure 5.28: The integral of each CAEN trace from the MPU for a subrun of SMELLIE data (using PQ440) plotted as a function of event GTID. The subrun is the same as in Fig. 5.27.

show any time dependent behaviour on the same scale as the heights. This indicates that the variation in heights is therefore not an artefact of the pulse shape changing or the baseline varying.

To explore the time dependency shown in the CAEN heights, the CAEN heights as a function of time (taking the first SMELLIE event of the subrun as 0 ns) were Fourier transformed to discover the underlying frequencies behind the periodic variations. This, for the subrun shown in Fig. 5.27, is shown in Fig 5.30. The frequency spectrum extends to 500 Hz as this is the Nyquist frequency when the sampling rate is 1 kHz, as it is in this subrun. There are clear peaks in this frequency spectrum. It is hypothesised that these could be due to racks of electronics on the deck being located close to the SMELLIE racks. However, despite this procedure being performed for many SMELLIE subruns and consultation with the electronics group, these frequencies peaks were not successfully matched to known causes.

Additionally, it is fundamentally important to understand the relationship between the height of the CAEN trace and the number of photons that enter the detector. If it is not true that these are correlated, then the MPU is not performing the role it was designed for. Fig. 5.31 shows the CAEN heights as a function of the cleaned NHits in the detector for the same subrun as Fig. 5.19 (i.e. a subrun with stable laser behaviour), while Fig. 5.32 shows the same variables for the subrun in Fig. 5.21 (i.e.
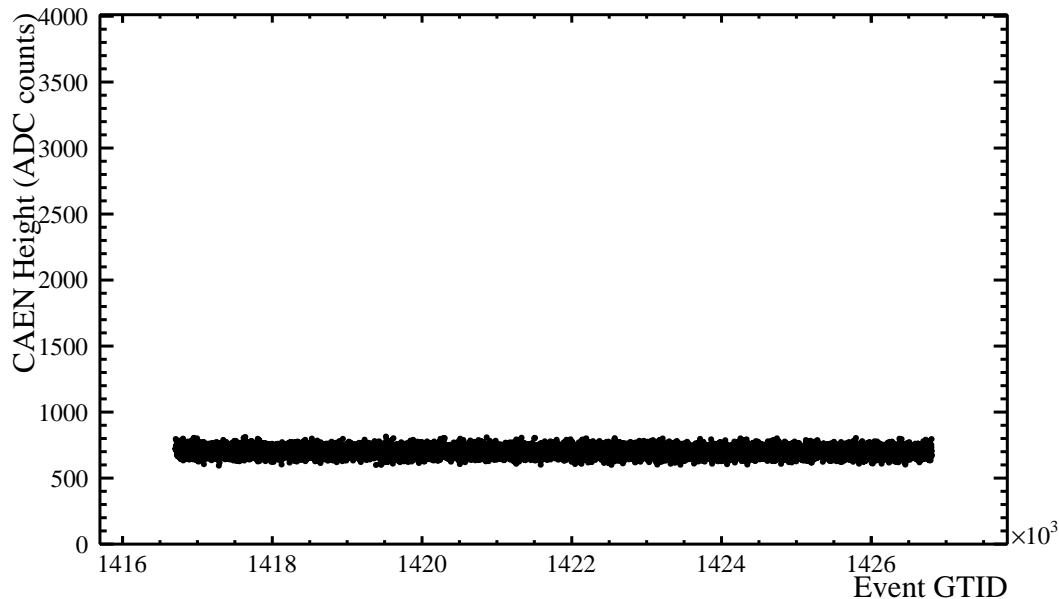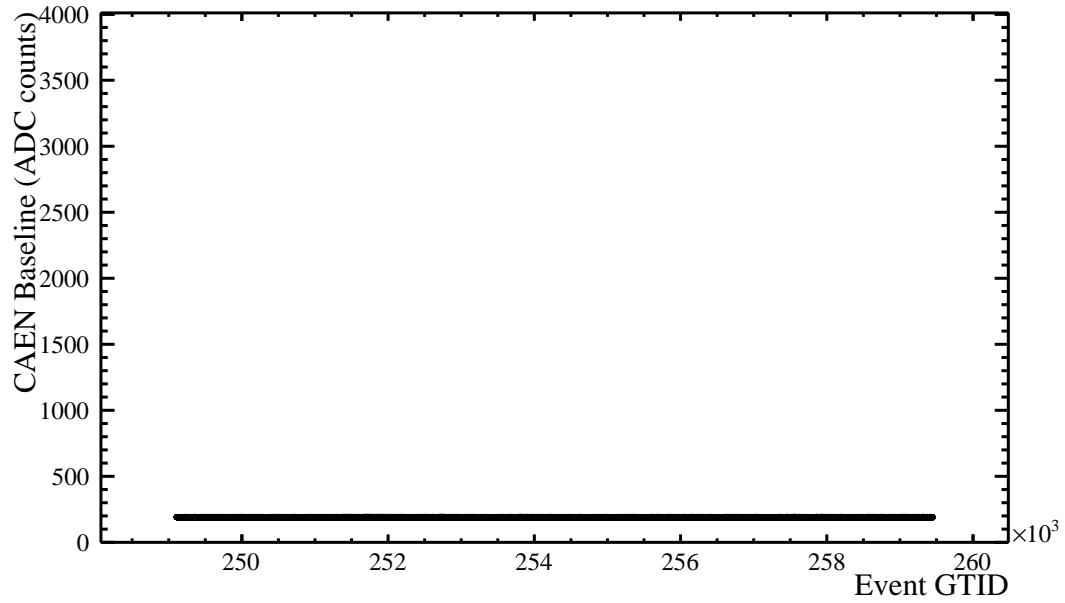
Figure 5.29: The baseline of each CAEN trace from the MPU for a subrun of SMELLIE data (using PQ440) plotted as a function of event GTID. The subrun is the same as in Fig. 5.27.
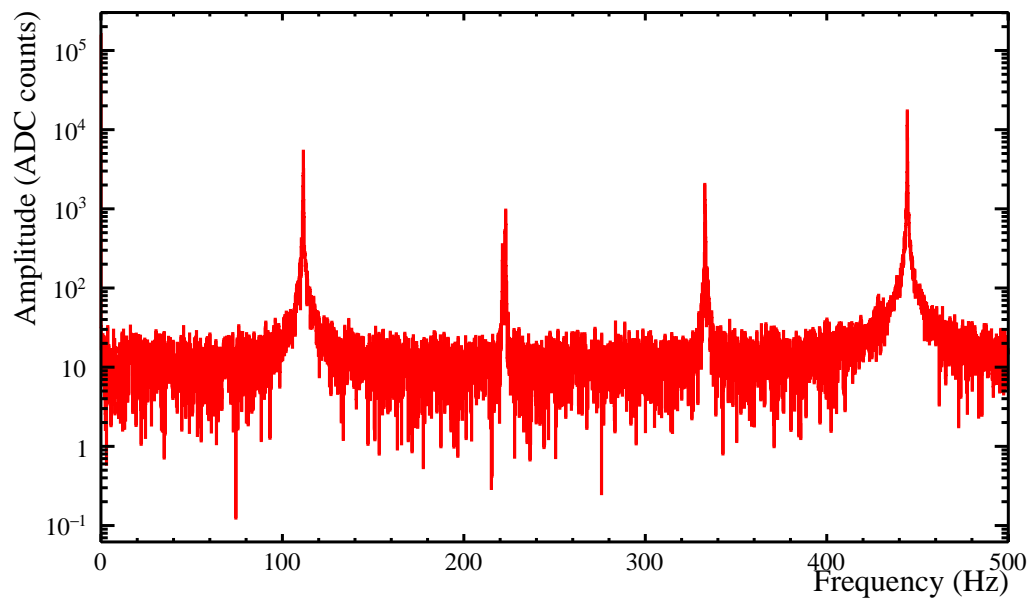
Figure 5.30: The Fourier transform of Fig. 5.27 into frequency space.
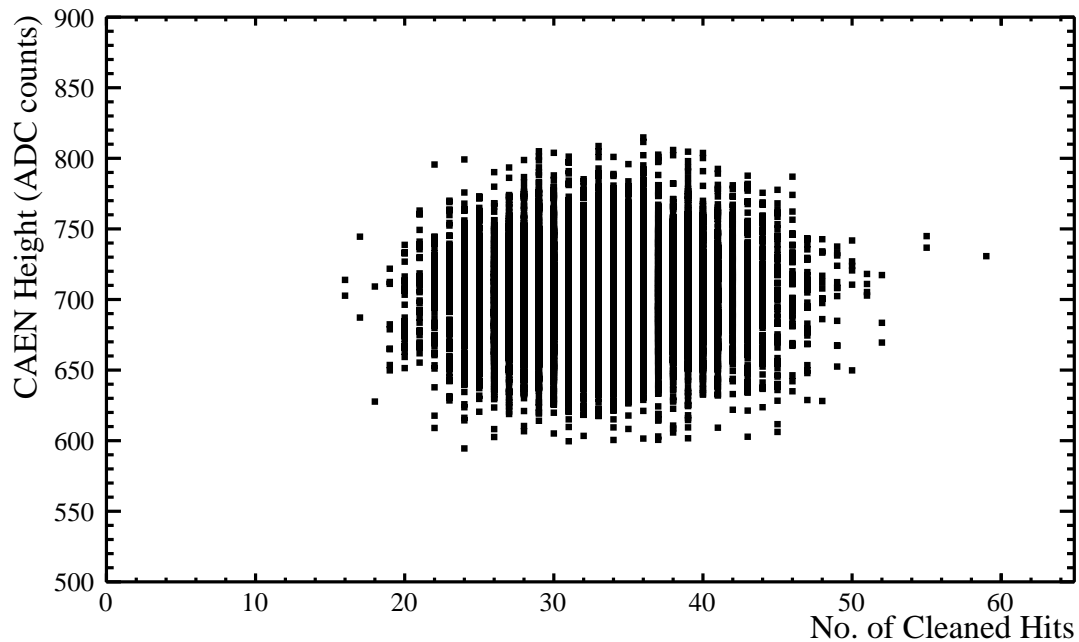
Figure 5.31: The height of CAEN traces for a subrun of SMELLIE data (using the superK at 495 nm) plotted against the cleaned NHits detected for each event. This plot is made using the same subrun as Fig. 5.19 and shows stable laser running.

a subrun with unstable laser behaviour). Cleaned NHits (hits which have passed a predefined set of data cleaning cuts designed to remove instrumental artifacts) are used as an approximation of photons entering the detector. Obviously, this is not strictly true, especially in the case of high shot energies. However, the multi-hit correction, as used in this work, only calculates the average p.e. in a given PMT over a subrun, which is not useful for comparing fluctuations to CAEN height on an event to event basis.

As can be seen in these plots, in the case of unstable laser running, the cleaned NHits are correlated to the CAEN height. This, as expected, is not a linear relationship. However, it is evident that an increase in the CAEN height does result in a larger number of photons in the detector. (A correlation is not expected to be seen for a subrun of stable laser running as every laser shot is at the same intensity and therefore the NHits and CAEN heights are centred on a single value.)

When choosing subruns for the scattering analysis, an effort was made to ensure that subruns were chosen with either stable laser behaviour or as low a level of instability as possible. Nonetheless, the impact that variations in the shot energy have on the result of the scaling factor has been investigated more fully using the CAEN trace variations in Section 7.4.6.
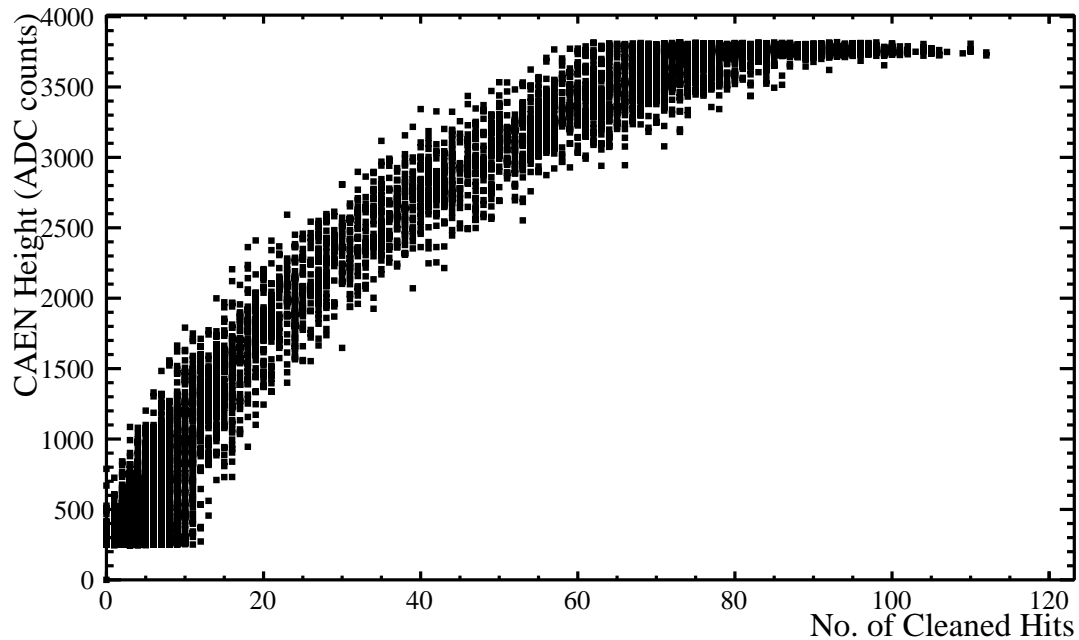
Figure 5.32: The height of CAEN traces for a subrun of SMELLIE data (using PQ440) plotted against the cleaned NHits detected for each event. This plot is made using the same subrun as Fig. 5.21 and shows laser instability.

5.6 Tesseract

When firing SMELLIE, a laser, wavelength, shot energy (or intensity) setting, MPU gain and fibre have to be specified. The combination of these variables are chosen with the aim of producing a desired number of p.e. in the detector and usuable CAEN heights.

The relationship between these settings and the number of p.e. detected is not simple. To complicate the situation further, the number of possible combinations of settings is extremely large. 15 fibres $\times$ 33 wavelengths/lasers $\times$ $\sim$100 shot energies $\times$ $\sim$100 gains = 4,950,000 combinations. At 10000 shots fired per combination, at a rate of 1 kHz, and without any deadtime, this would take over 1.5 years of live detector time. Consequently, to explore the relationship by taking data for each combination of settings is not possible.

Therefore, the SMELLIE tesseract was developed. The tesseract (named for being a 4D hypercube where the axes are fibre, wavelength/laser, shot energy and gain) is a database of settings at which data was taken, together with interpolation and extrapolation routines. The aim is to be able to use a limited number of datasets in order to predict settings which achieve a desired number of p.e. and CAEN height for

each shot.

To build the tesseract, runs were taken that formed 1D slices of the 4D space. One of these runs, for example, would be 15 subruns long, with each subrun using a different fibre while keeping all other settings constant. A run like this is referred to as a fibre scan. Similar scans through wavelengths/lasers, gains and shot energies, as well as intensity scans for several fibres, were taken. From each subrun in every one of these runs certain variables of interest, discussed below, were extracted by the author. Using these variables, J. Lidgard created interpolation and extrapolation routines. Testing was performed by using the routines to predict a series of settings, firing SMELLIE subruns at those settings and calculating how far away from the desired p.e. number and CAEN height the shots in the subrun were. Using the extra data from these tests, the routines were improved and the process was repeated. For further detail on the development of the tesseract, and, in particular, the interpolation and extrapolation techniques used, see [103].

There is interest in three regions of shot energy in SMELLIE data:

- Low - Used for characterising beam profiles. A region where multi-hits are not a large correction.

- Medium - Used for the scattering analysis. A region that is multi-hit correctable (see section 5.2), therefore providing greater statistics.

- High - Possible to gather enough statistics at high scattering angles to analyse the angular distribution of scattering. In the direction of the beam, multi-hit corrections will not be possible and therefore a charge based analysis must be done.

In order to build the tesseract, detector quantities have to be extracted from each subrun of data. These quantities have multiple purposes. They must be used to define each of the above regions and therefore be used to test whether the subrun achieved the desired shot energy. The quantities extracted from the subrun must also be used in the extrapolation and interpolation routines of the tesseract. In addition, it is desired that sufficient information is gathered to judge whether the subrun could be used for the scattering analysis. While this obviously requires the desired shot energy regime to be achieved, it also means that information such as the variation of the CAEN heights during the subrun should be recorded to judge the stability of the laser (as discussed in Section 5.5). Therefore, taking this all into consideration, the detector quantities of interest for each subrun are as follows:

- Maximum occupancy - The maximum occupancy in any PMT during a subrun. This is a measure of whether data is multi-hit correctable. As previously discussed, once an occupancy of one is reached in a PMT, the p.e. in that PMT cannot be statistically calculated from the NHits on it over the course of the subrun.

- Average top QHS - The average QHS per shot for each one of the seven PMTs which have the highest QHS summed over the subrun. This quantity is designed the replace the maximum occupancy measurement in the regime in which the multi-hit correction cannot be applied. However, simply taking the maximum QHS of any PMT, as is done for the maximum occupancy, does not give good results. There are PMTs for which the QHS measured is saturated in any given event with no clear cause, which would result in most subruns being deemed unusable. Therefore, the average of a number of PMTs is used to give a clearer indication of whether QHS is saturated due to the SMELLIE pulse, meaning that p.e. calculations using QHS can no longer be performed. This measure also ensures safe operation of the lasers with respect to the detector, as saturating QHS in many PMTs belonging to the same electronics crate with a fast repetition rate could run the risk of either PMT or electronics damage.

- Average p.e. from occupancy - The average p.e. estimated to be in the detector per shot calculated from multi-hit corrected NHits.

- Average p.e. from QHS - The average p.e. per shot estimated to be in the detector, calculated from QHS on a PMT by PMT basis by dividing the QHS measured by the HHP of the PMT.

- Maximum CAEN height - The maximum CAEN height of any SMELLIE event during the subrun.

- Average CAEN height - The average CAEN height of all the SMELLIE events in the subrun.

- RMS CAEN height - The RMS (root mean squared) CAEN height of all of the SMELLIE events in the subrun.

Before these quantities are calculated for a subrun, PMTs showing irregular behaviour are excluded using routines described in Section 5.6.1.
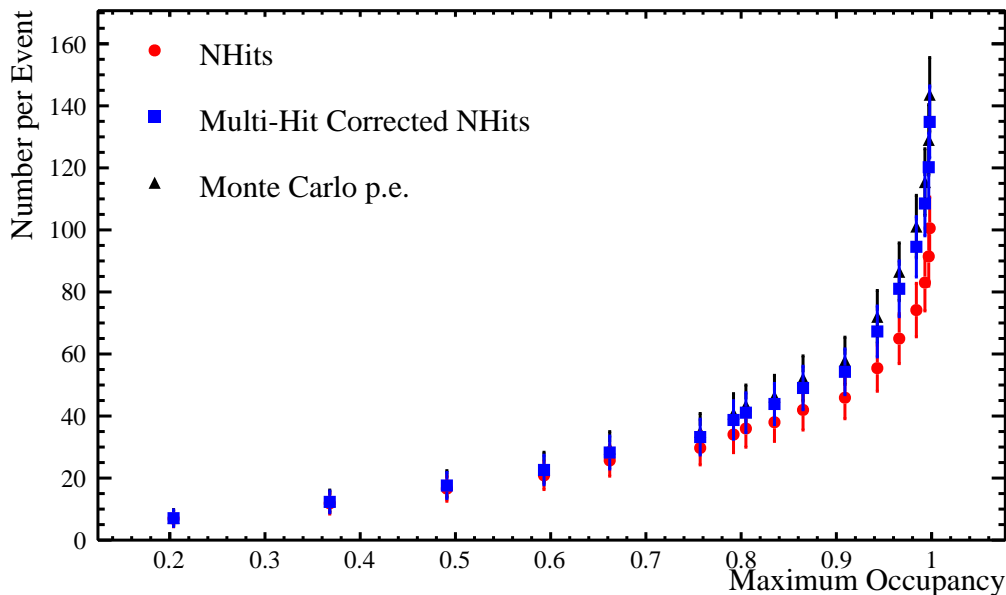
Figure 5.33: NHits (red circles), multi-hit corrected NHits (blue squares) and MC p.e. (black triangles) per event for a simulated SMELLIE subrun plotted as a function of the maximum occupancy for any PMT during the subrun. FS007 was used with the supercontinuum laser at 495 nm. The simulation used run level conditions for run 114018. The error bars are Poissonian.

The three shot energy regions must be defined in terms of the above detector quantities. Low shot energy is defined as a maximum occupancy of 0.35, medium a maximum occupancy of 0.8 and high an average top QHS of 1200 ADC counts (approximately 3 tenths of the maximum QHS). The two occupancy definitions arise from Fig. 5.33 which compares the NHits, multi-hit corrected NHits and Monte Carlo p.e. per event in a series of simulated subruns of FS007 at 495 nm using the superK.

The aim of the low occupancy regime is to provide data to create accurate beam profiles, as discussed in Chapter 6. Ideally, for this purpose, the PMT occupancy for all PMTs would be in the single p.e. regime. As seen in Fig. 5.33, in subruns with a maximum occupancy under approximately 0.5, the NHits, multi-hit corrected NHits and Monte Carlo p.e. are all indistinguishable quantities. Therefore, a maximum occupancy of 0.35 was chosen as the desired value for low shot energy subruns, in order to provide a margin for error. This is required as SMELLIE runs with identical settings do not always produce identical results; laser intensities can vary as well as the detector configuration.

A similar logic is applied to the medium shot energy region. This is defined as aiming for a maximum occupancy of 0.8. This is because, as shown in Fig. 5.33, at a maximum occupancy of 0.8, the multi-hit corrected NHits are completely consistent

Shot Energy Region	Laser	Wavelength (nm)	Fibre	Shot Energy	Gain
Low	PQ375	375	FS007	840	0.477
Medium	PQ375	375	FS007	851	0.395
High	PQ375	375	FS007	1000	0.334
Low	PQ375	375	FS107	843	0.453
Medium	PQ375	375	FS107	854	0.39
High	PQ375	375	FS107	1000	0.334
Low	PQ495	495	FS055	294	0.751
Medium	PQ495	495	FS055	433	0.683
High	PQ495	495	FS055	729	0.547
Low	superK	495	FS055	245	0.687
Medium	superK	495	FS055	256	0.623
High	superK	495	FS055	282	0.512

Table 5.1: The shot energy and gain settings produced by the tesseract for a small, illustrative, selection of fibre, wavelength/laser, shot energy combinations. These settings were used in runs 114018 and 114030. Subruns from run 114018 were used in the scattering analysis.

with the Monte Carlo p.e., whereas these are both systematically larger values than the NHits. This demonstrates that the multi-hit correction is valid to use at these occupancy values. This is also true for occupancy values above 0.9; as with the low shot energy region a margin for error is left. A safety margin against reaching an occupancy of 1 is particularly important for the medium shot energy region because if this occurs in any PMT during a subrun, that subrun can no longer be used in the standard scattering analysis.

The desired average top QHS used to define the high shot energy region was simply chosen such that subruns have some PMTs with saturated occupancy while the QHS of the PMTs is not saturated. This means that there are greater numbers of hits at high scattering angles, allowing further insight into the angular distribution of the scattering and hence the type of scattering occurring. Keeping the QHS unsaturated means that the number of p.e. in each PMT can be estimated using the HHPs. However, this high shot energy data is not analysed as part of this work.

The tesseract was successfully used to take SMELLIE data at shot energies useful for both beam profiling and performing the scattering analysis. As an example of some of the settings resulting from the tesseract, Table 5.1 shows a selection of the settings used for run 114018 (a run used heavily in the water phase scattering analysis) and run 114030 (for operational reasons high shot energy data was taken in a separate run to low or medium shot energy data). The shot energy setting is on a scale of $0 - 1000$ and the gain $0 - 1$.

While, in general, the tesseract fulfilled its purpose, the settings shown in Table 5.1 do highlight some of the issues that arose during the commissioning. Attenuation adjustments have already been discussed in Section 5.5 and here it can be seen that the high intensity settings for the PQ375 laser are at a maximum of 1000. As mentioned before, the attenuator was not adjusted for PQ375, meaning that the high shot energy subruns taken at 375 nm never reached the intended average top QHS value. Further work on the attenuators was intended to happen during scintillator fill. However, restrictions due to Covid-19 mean that this work has not occurred.

5.6.1 Data Cleaning for Commissioning

In the creation of the tesseract, PMTs showing irregular behaviour are excluded. The SNO+ code base now includes a tool called Channel Software Status (CSS), primarily for the purpose of finding PMTs which should be excluded from analyses due to irregular behaviour [108]. However, as building the tesseract occurred so early in detector commissioning, CSS was not functional at the time. Therefore the author had to define irregular behaviour and identify it. The PMTs that need removing are those commonly called screamers, otherwise known as high occupancy PMTs. These are defined in CSS by calculating the total number of hits per PMT during PGT events as a proportion of the total number of hits during PGT events. The mean of this is found for all PMTs and those PMTs whose rate is greater than three RMS deviations above the mean are selected as high occupancy.

However, as already mentioned, this definition was decided on after the building of the tesseract. For tesseract purposes, the crucial reason for the removal of screamers is to ensure they do not cause erroneous maximum occupancy measurements. Therefore, they are identified by finding any PMTs with an $\alpha > 0.1$ rads with one of the ten highest summed NHits over the subrun in question. (PMTs with an $\alpha > 0.1$ rads are PMTs are outside the central beamspot, hence high occupancies are not expected.) Any PMT identified as a screamer during a subrun of SMELLIE data is excluded from the calculations of the above detector quantities for that subrun.

Another calculation that must be performed for each subrun is the subtraction of noise. This would now be done using the noise rate tables in RATDB which store the rate of noise according to PGT events for each PMT for each run. However, as before, this table was not available when the tesseract was built and therefore the calculations had to be performed manually. For the tesseract, the noise rates are calculated using PGT events in the run proceeding the SMELLIE run. (If the state of the detector dramatically changes during the SMELLIE run, the run after the SMELLIE run is

used for subruns post detector change.) The noise rate per PMT is not calculated, as the subtraction is only applied to the average p.e. calculations. Therefore, for the average p.e. from occupancy, the average NHits per PGT event in the run is calculated and subtracted. NHits can be used as noise hits are too low occupancy for a multi-hit correction to be relevant. For the noise subtraction on the average p.e. from QHS, a QHS value needs to be calculated for the noise hits. Therefore, for each PMT hit in the PGT events, the QHS is converted into p.e. and the total p.e. from QHS measurements is calculated. The average of this across all PGT events is taken to provide an alternative measure of the noise rate, to be consistent with the p.e. calculation being performed.

By the time of the scattering analysis, described in Chapter 7, noise rate tables had been created for all runs, including those from commissioning. Therefore a noise correction does not need to be manually applied in the scattering analysis - using run level settings in simulation means that the noise rate tables are used to accurately simulate noise and screamers.

5.7 Data Runs

Over the course of this work, approximately 100 SMELLIE runs were performed, each often consisting of several hundred subruns, each with a different combination of settings. Therefore, it is infeasible to summarise all of the SMELLIE data taken here. Much of it was simply used for commissioning, with various aspects of that described in this chapter. In particular, creating the tesseract involved iterations of predicting settings, testing runs, evaluating, interpolating new predicted settings and so forth. This process generated a lot of runs that could be utilised in the future, but have so far only been used to inform the next iteration of the tesseract.

There are two main sets of data referenced in this work. Both are runs taken during the iterative process of building the tesseract. One is the first set of runs taken in this process. These runs are only used in the commissioning of SMELLIE. They provide a larger dataset (including attempts at high, medium and low shot energies for all fibres at all wavelengths) to do detailed commissioning work with than any of the runs which preceded them. However, as this set of runs was the first in the tesseract building process, many of the settings were not optimal and so some subruns were at too high or low shot energy to be fully utilised. This set of runs was taken on 18/11/2017 and covers run numbers 106845 - 106855.

The second set of data is the final iteration of the tesseract taken in the water phase. (Further SMELLIE runs were taken after this, but for the purpose of preparing for SMELLIE operations in scintillator phase, a topic not covered in this work other than a brief discussion in Chapter 8.) Therefore, this set of runs represents the cumulation of all the knowledge acquired regarding running SMELLIE in water phase and are typically the closest to the desired shot energies. It was used both for beam profiling (see Chapter 6) and for the scattering analysis (see Chapter 7). It was additionally used for some aspects of commissioning that occurred later than the initial tesseract work. This set of runs was taken approximately seven months after the first set on 20/06/2018 and covers run numbers 114018 - 114039.

All of these runs were taken at a pulse rate of 1 kHz and with 10000 shots per subrun. 1 kHz was chosen as a rate which allowed data taking to happen as rapidly as possible while not causing issues in the detector electronics (high intensities at higher rates can cause FIFO overflow errors to occur in the XL3s) and not running any risk of causing damage to PMTs.

Chapter 6

Simulating SMELLIE

6.1 SMELLIE in Simulation

The SMELLIE system is simulated in RAT only through the injection of photons into the detector at the positions of the fibre ends. These simulations are governed by the ELLIE generator, which is shared with TELLIE and AMELLIE. It uses database tables both for fibre and for the source properties. For SMELLIE, the sources are the five lasers (PQ375, PQ405, PQ445, PQ495 and superK). Among the source tables are tables for the emission time distributions and wavelength spectra of each laser. These are taken from the laser specifications and shown in Figs. 4.4 and 4.3, respectively. These tables can be swapped out for alternative distributions for testing purposes.

The fibre tables contain the fibre position and direction information as well as the angular emission distribution for that fibre. Again, these tables can be swapped out in the simulation macro and alternate position and direction information can be supplied. The directions and positions of the fibres are discussed in Section 4.1.2. In this work, the angular distribution tables are also referred to as beam profiles and the creation of these is referred to as beam profiling.

At present, the angular emission distribution is considered to be independent of wavelength. The major concern involved with this assumption - that there may be diffraction patterns in the beam profiles - has been shown, in Section 5.4, not to be warranted. The timing and angular distributions are also assumed to be independent of each other. While this is not strictly true due to modal dispersion in the optical fibres, the fibres used in SMELLIE have extremely low modal dispersion, justifying this assumption [53].

When a simulation is run, photons are thrown from the fibre position according to the angular, timing and wavelength distribution tables that are associated with the fibre and source that have been specified in the macro. There are two generator

methods within the ELLIE generator for creating the angular distribution specified. These will be discussed in the following sections, as will the creation of the angular distribution tables for each generator method.

6.1.1 1D Generator

The 1D ELLIE generator can be used for all three of the ELLIE subsystems. The first version of this generator was written by J. Wilson in 2008 [109] and, until this work, has been the sole method of generating ELLIE beams.

The 1D generator is called so because it creates a beam which is azimuthally symmetric around the fibre direction and only allows variations of the emission probability with the polar angle, α, as defined in Section 5.1.

The angular distribution table used for this generator has two fields that are used as an input to the generator method: a list of α angles and a list of relative probabilities at those angles. The α angles must be monotonically increasing and the number of entries in the two lists must be the same. The α angle list can be in either radians or degrees as there is a separate field in which the angular units are specified. If degrees are used, the list will be converted into radians. The first and last entries in the angle list are also checked to make sure they are both positive.

Photons will be randomly thrown in α according to the specified probabilities, with linear interpolation applied between the α points. The method for throwing α angles in the 1D generator is also used for throwing the wavelength and time of each photon, using the wavelength and timing distribution tables of the specified source. Independently, the ϕ angle is randomly sampled from a flat distribution between $0 - 2\pi$ rads.

The definition of ϕ in the 1D generator had to be changed by the author in order to make it consistent with the definition presented in Section 5.1. Previously, a consistent definition of ϕ across fibres was not important.

It should be noted that the 1D generator applies a solid angle correction to the relative probabilities provided so that if a distribution is provided that is flat in α, the thrown distribution will be flat in $\cos(\alpha)$.

Certain angular distributions have been implemented in the generator (for testing purposes) that do not involve writing a new angular distribution table. These are a pencil beam (i.e. a beam with no angular dispersion) and an isotropic beam (flat in $\cos(\alpha)$) up to an α of 90 degrees.

6.1.2 2D Generator

The 2D generator was written by the author and is currently only used by SMELLIE, though AMELLIE plans to adopt it in the near future. As discussed in Section 6.2.1, after commissioning runs were taken with SMELLIE, it became apparent that the angular distribution of photons emitted from the fibre in data did not match simulation. Considering that using profiles which are symmetric in azimuthal angle has always been a simplification, the opportunity was taken to write a generator that allowed variation not only in polar angle, but also in azimuthal angle.

The 2D generator takes a different angular distribution table as an input than the 1D generator. The two fields which define the 2D beam profile are a list of PMT IDs and a list of relative photon intensities at each of these PMTs. The creation of this table will be discussed below in Section 6.2.2. This second list is normalised such that the maximum value in it is equal to one.

Firstly, like the 1D generator, the 2D generator reads the fibre ID from the macro and loads the position and direction information for that fibre. The PMT IDs are then used to access the PMT positions and, for each of the PMTs listed, the positions are converted into polar coordinates with respect to the fibre direction (as described in Section 5.1).

The generator uses an acceptance-rejection Monte Carlo method. A test pair of polar coordinates are thrown, uniformly between zero and 2π for ϕ and, for α, uniformly between zero and the maximum α of any PMT in the PMT ID list. From the test coordinates, the point of intersection (in detector coordinates) of a photon fired using the test coordinates and the PSUP is calculated. This is called the hit location. The three nearest PMTs to the hit location are identified and, using the relative intensities at each PMT, an intensity for the hit location is interpolated. This interpolation takes the form of an inverse cubic distance weighted average. This average is calculated as

$$I_{int} = \frac{\frac{I_a}{D_a^3} + \frac{I_b}{D_b^3} + \frac{I_c}{D_c^3}}{\frac{1}{D_a^3} + \frac{1}{D_b^3} + \frac{1}{D_c^3}}, \tag{6.1}$$

in which I_{int} is the interpolated intensity at the hit location, I_i is the intensity at PMT i and D_i is the straight line distance to PMT i from the hit location.

This interpolated intensity is multiplied by $\sin(\alpha)$ in order to correct for uniformly sampling in α rather than flat in $\cos(\alpha)$ (causing the area around the beamspot centre to be sampled with increased frequency).

Next, a test intensity is thrown from a uniform distribution between zero and one. If this intensity is smaller than the interpolated intensity, the test coordinates are accepted and a photon is fired in the simulation. If the test intensity is larger than the interpolated intensity, the test coordinate pair is rejected and another is thrown. This process is repeated until the requested number of photons have been fired.

This process is significantly slower than the 1D generator due to the number of test coordinates thrown at the edges of the beam where the intensity is low. For each, the three nearest PMTs to the hit location must be found and the intensity interpolated. Therefore, a Gaussian envelope in the α coordinate is implemented to increase the speed of the generator. This means that rather than throwing both α and ϕ flatly, the test α is thrown according to a Gaussian distribution centred on zero with an amplitude of 1. Additionally, the test intensity is now thrown between zero and the value of the Gaussian function at the test α. Therefore, test coordinates at higher values of α are thrown less frequently and they are less likely to be rejected.

Ideally, the distribution used for the envelope should be a close match to the α distribution of the profile, therefore minimising the number of failed test coordinates thrown. However, it has to have an amplitude greater than the profile for every α. If this is not true, an error will be thrown by the generator and the simulation will be aborted.

Due to the general shape of the profiles, a Gaussian distribution was chosen for the envelope. As previously stated, the Gaussian is centred around $\alpha = 0$ rads (the beamspot centre) and it, by default, has a sigma of 0.5 rads. This sigma can be altered in the RAT macro (as can the other parameters of the Gaussian). 0.5 rads ensures that the amplitude of the profile is less than the amplitude of the Gaussian, even in the tails, for all of the beam profiles. This is a larger sigma than might seem necessary; however, it is needed to correctly simulate the rings present in several of the profiles. It would be possible to tune this parameter to the profile for each fibre individually, hence speeding up the simulation for fibres without ring features. However, this strategy was not pursued in this work.

Despite the speed increase associated with using a Gaussian envelope, the 2D generator is still significantly slower than the 1D generator. For example, to simulate 100 photons at 495 nm using FS007 takes 20 times longer when using the 2D generator and full 2D profile, compared to the existing 1D profile and generator. This, considering the amount of simulation necessary to run the full scattering analysis, is not a trivial consideration. However, having applied some code optimisation to

the generator to result in this factor of 20, it was decided that this was a workable simulation speed. Ideally, work would be continued to optimise the generator further.

Although inverse cubic distance weighting is used in the final version of the 2D generator, other interpolation methods were tested: linear, quadratic, quartic, decic and exponential inverse distance weightings. To conduct these tests, PMT by PMT comparisons were made between data and simulation (created using the 2D profile extracted from data) as well as comparisons between two simulations - one using a top hat profile and the 1D generator and one using a top hat profile and the 2D generator. To make these comparisons, a χ^2 variable is defined for each PMT as

$$\chi^2 = \frac{(\mathrm{p.e.}_a - \mathrm{p.e.}_b)^2}{\sigma_a^2 + \sigma_b^2}, \tag{6.2}$$

in which p.e.$_i$ is the average p.e. detected by PMT i per shot and σ_i is the statistical error associated with that p.e. value. a and b refer either to data and simulation, or to simulation with the 1D profile and simulation with the 2D profile. A histogram is then created of these χ^2s. This histogram should follow a χ^2 distribution, if the difference between the two data sets is statistical. Therefore, the histogram for each interpolation method is compared to a χ^2 distribution. This showed that exponential interpolation is a worse method than polynomial interpolation. Also, generally, the higher the exponent, the better the match. However, the difference between cubic and quartic (or decic) interpolation is negligible. Therefore, cubic interpolation is used for the 2D generator. Further exploration of the match between data and simulation created using the 2D beam profiles will be explored in Section 6.2.2.

One final consideration when creating the 2D generator was the definition of the polar angle, α, as defined in Section 5.1. α uses the vector from the fibre to the PMT in question in its definition. This does not take into account refraction through the AV - it assumes that the detector is empty of anything that could alter the photon direction after it has been fired and that the hit pattern of the PMTs is representative of the distribution of light leaving the fibre. Using the `LightPathCalculator` class, it is possible to take the fibre position and the PMT position and recreate a photon path between them that does account for refraction through the AV. From this path, the initial direction vector of the photon can be extracted. Therefore, another angle, β, is defined almost identically to α, but rather than using the vector from the fibre position to the PMT, the initial direction vector of the photon is used.

Simulations were created and compared using α and β for FS055 at 495 nm with the same 2D profile. No systematic difference was found between the two simulations,

even at higher values of α, where incident angles of photons on the AV (and there-
fore refraction effects) will be larger. While it is true that FS055 is not a 20° fibre
and therefore not the most extreme case, it is a 10° fibre. Estimates of the change
in hit location that would be expected for photons fired along the fibre direction
were found to be on the order of mm (negligible compared to the PMT size). Such
estimates combined with the complete lack of any statistically significant systematic
change in the simulation created with β rather than α was taken as sufficient evidence
that this slight inaccurate use of α when generating beams is negligible. As using
`LightPathCalculator` to find the initial photon vector for each PMT increases the
computation time of the generator, using α is deemed sufficient.

6.2   Creating Beam Profiles

For both the 1D and 2D generators, an angular distribution input table is required
for each fibre, to specify the relative probability of a photon being emitted in a given
direction, in the required dimensions, with the required fields. The format of these
tables has been discussed above. Next, the creation of these beam profiles is discussed
and the results of the combined beam profile and generator are compared to data.

6.2.1   1D Profiles

The tables forming the 1D profiles have been briefly discussed in terms of the fields
required in Section 6.1.1. To briefly recap, there is a list of monotonically increasing
α angles and a matching list of relative probability values at each respective α.

The 1D profiles currently in RAT were created by S. Langrock and the process of
creating them is discussed in [104]. In short, three fibres which were installed at that
time and had sufficient data taken (FS037, FS237 and FS155) to have their angular
distribution measured in air-fill data. For the remaining nine (excluding the 21 node
fibres) data was used from CCD camera images taken at Oxford. For the 21 node
fibres, an average profile based on the three air-fill profiles was created as an estimate
of a realistic profile.

Simulations created using the 1D profiles for FS055, FS007 and FS237 are shown
in Figs. 6.1 - 6.3. Also shown are the data subruns which the simulations have been
based on. It is clear by eye from these plots that the simulations are not good fits to
the data. In these plots the χ^2 per bin between data and simulation is also shown.
This is below the main plot and is calculated by taking the square of the difference
between the data and the simulation and dividing it by the quadratic combination
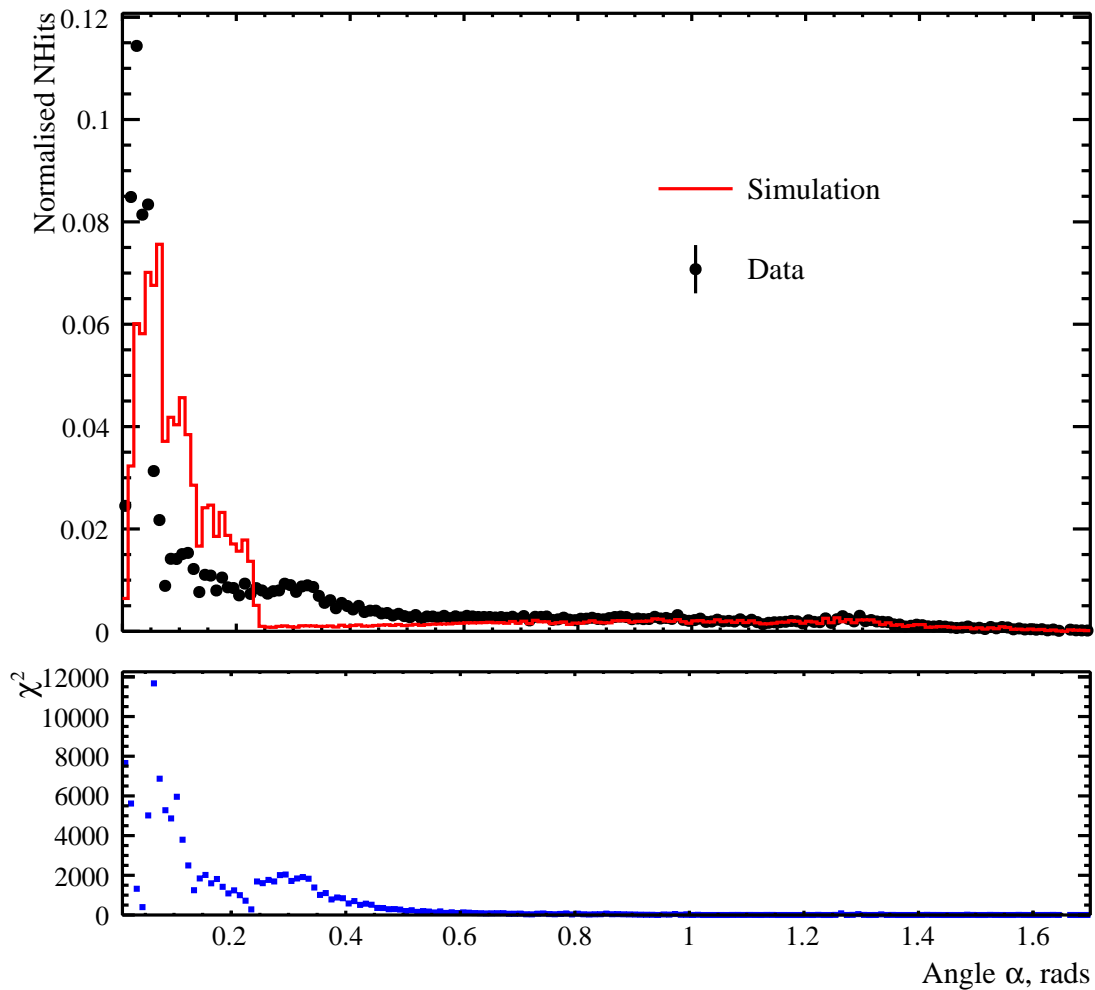
Figure 6.1: Simulation created using the 1D beam profile of FS055 at 495 nm is shown in red, compared to the a data subrun in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 653.22.

of the two associated errors. The average of this χ^2 is also calculated, across all the bins, and is referred to as the χ^2_{DOF}. None of the χ^2_{DOF} for these plots are below 650, indicating an extremely poor match.

One characteristic of these plots is, in both data and simulation, that the content of first bin is significantly lower than the second and that the peak of the distribution is a few bins away from zero. This is a geometrical effect caused by using a constant bin width in α. In essence, each bin represents a ring of PMTs within a polar angle range. This means that rings of constant α width at larger values of α are wider in 3D space than at smaller values of α and therefore, there are simply very few PMTs in the first couple of bins in the plots. It is possible to correct the plots for this effect
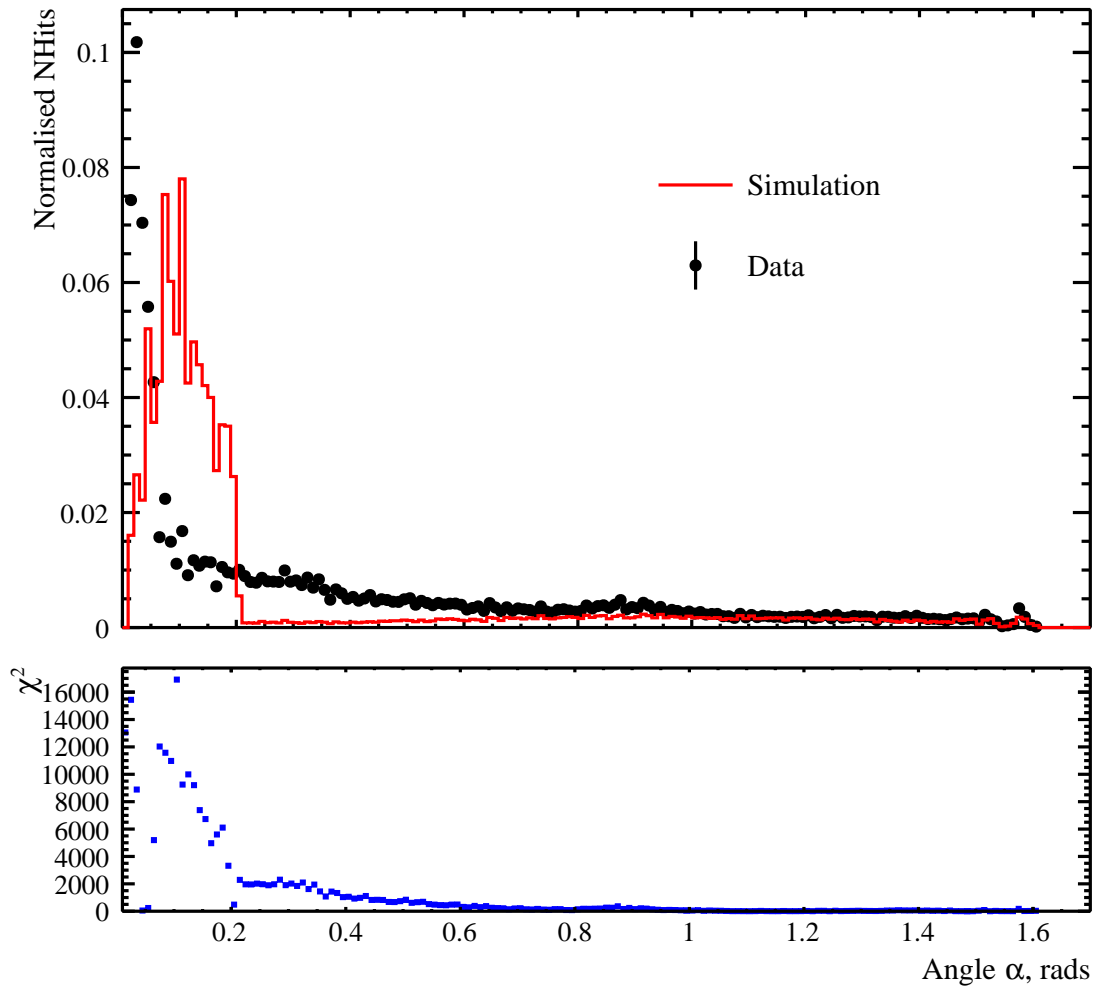
Figure 6.2: Simulation created using the 1D beam profile of FS007 at 495 nm is shown in red, compared to the a data subrun in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 1264.04.
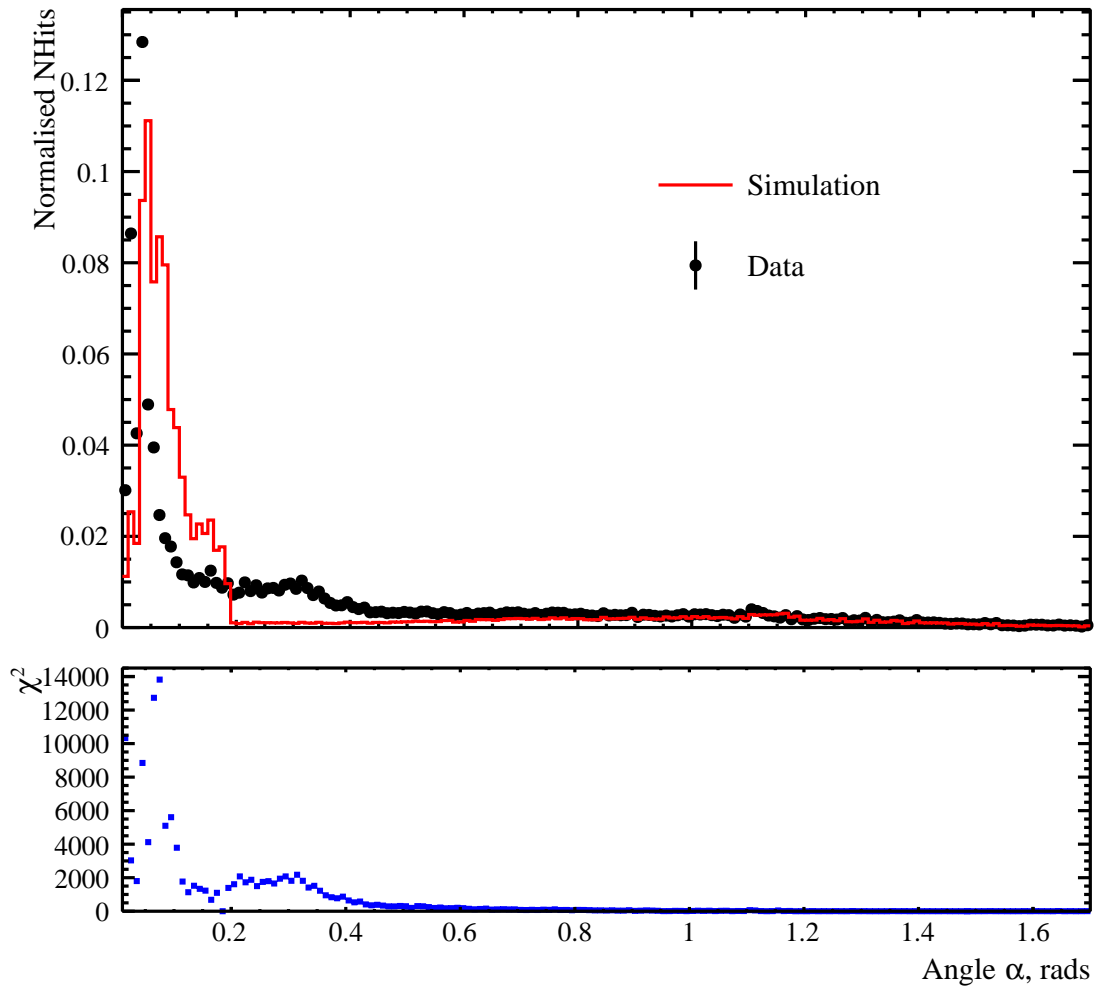
Figure 6.3: Simulation created using the 1D beam profile of FS237 at 495 nm is shown in red, compared to the a data subrun in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 724.92.

by dividing each bin by the number of PMTs within it. However, as both simulation and data suffer from the effect identically, this is not done here.

Although it already seems clear that these 1D profiles are not a good match to data, the scattering analysis, as described in Chapter 7, was run once using the 1D FS055 profile. Doing this confirmed the hypothesis that the simulation and data are simply not well matched enough when using the existing 1D profiles for the result of such an analysis to be meaningful. This is due to the analysis being highly dependent on Monte Carlo - the difference in angular hit distribution meant that it appeared that many more photons had undergone scattering in data, when in reality, these photons were simply part of a wider beam than modelled in simulation.

2D profiles were already desired, for all of the advantages previously discussed. However, after the 1D profiles had been tested by using them for the scattering analysis, it became crucial that the 2D generator and beam profiles were created before an accurate measurement of the scattering length could be made.

6.2.2 2D Profiles

The aim of the 2D profiles is to extract data from un-scattered photons near the centre of the beam and model the angular distribution based on these observations.

Similarly to the tables forming the 1D profiles, the 2D profiles each consist of two lists. One of those is a list of PMT IDs within a certain polar angle from the beam-spot centre and the other is a list of intensity values measured at each of the PMTs. The translation of PMT IDs into coordinates and the interpolation to points between PMTs is done completely within the generator itself, as has been discussed in Section 6.1.2. While using a PMT ID list makes creating test profiles more challenging, the 2D profiles are designed to be created from detector data, which means that PMT positions are the natural coordinates to use. Using a polar grid of α and ϕ coordinates as the input was considered, but this would require interpolation of the intensity values in both the creation of the beam profile, then again in the generator. This was considered to simply be more complicated than necessary and therefore introduce more potential for error.

In order to create the 2D profiles, water-fill detector data was used. The low shot energy data set was taken specifically for the purpose of beam profiling. These were subruns taken in which the maximum occupancy in the detector was desired to be 0.35, meaning that the number of multi-hits should be small and certainly statistically correctable and should not bias the creation of the profiles. The importance of using low intensity data is discussed further in [104].

For the beam profiles created in this work, run 114018 was used. This run comes from one of the complete tesseract data sets taken, which means that low shot energy data was taken for all fibres. It is desirable for all the subruns to come from a single run (in which no HV trips or other such behaviour occurs) in order to ensure that detector conditions are the same for all the beam profiles. The wavelength of all of the subruns used is 495 nm. This was chosen mainly because the majority of the commissioning so far had occurred at 495 nm and therefore it was felt that the SMELLIE data taken at this wavelength was well understood.

Extracting the beam profile from data, however, is more complicated than simply using the hit distribution shown across the PMTs. This distribution contains artefacts of the detector geometry, such as such shadows, PMT efficiencies and concentrator angular acceptances, among others. Therefore, the distribution does not represent the distribution of photons being fired from the fibre and must be corrected for these detector effects.

While investigations were made into correcting each of these effects individually, it was realised that they could all be corrected together if the detector simulation was used to take account of them. To do this, for each subrun, an almost identical subrun is simulated. However, rather than a realistic SMELLIE beam, the angular distribution of photons fired from the fibre is isotropic (up to 90° away from the fibre direction), at a shot energy which ensures all the occupancies are multi-hit correctable. For each PMT, this occupancy is converted into the total number of p.e. in the PMT during the subrun. This is done for the data subrun as well. Then, the ratio of data to simulation for each PMT is calculated. (Prior to calculating the ratio, the p.e. in each PMT is divided by the maximum p.e. in any PMT for both data and simulation, in order to normalise out any difference in the number of shots.) This ratio is the 'intensity' variable in the 2D profile.

Using an isotropic beam for the simulation ensures that everything understood about the detector is taken into account: reflections, PMT efficiencies, shadows, concentrator angular acceptances, etc. It also means that any deviation from this isotropy in data that does not exist in simulation (such as the shape of the beam itself) remains a feature in the ratio.

Rayleigh scattering is turned off in these isotropic simulations as the aim is to correct for geometric effects as seen by photons exiting the SMELLIE fibre. Scattering blurs the image of these features. In addition, it must be remembered that these beam profiles are being created in order to measure the scattering in the detector

and, therefore, making an assumption about the scattering when making the profiles could bias the analysis.

The stages of this process are shown in Fig. 6.4. Each of the plots is made for FS007. The axes are polar, using the coordinate system described in Section 5.1. Each coloured circle represents a PMT and the colour scale represents $\log_{10}$ of the normalised number of p.e. (for data and simulation) or $\log_{10}$ of the ratio of p.e. in data to simulation. Fig. 6.4a shows the polar plot of the data intensity distribution across the PMTs while Fig. 6.4b shows the same but for the simulation using an isotropic beam. The ratio of data to simulation is shown in Fig. 6.4c.

The gaps in these plots are either areas in which PMTs do not exist or are areas of PMTs that were not on HV during run 114018. The isotropic simulations were created using run level conditions, meaning that any PMTs not at HV in the detector are not at HV in the simulation, either. For some fibres, there are patches of PMTs turned off very close to the beamspot centre. This is acceptable because the scattering analysis in this work is done using data either from run 114018 or from runs taken within hours of it and with an identical detector state.

The equivalents of Fig. 6.4c (i.e. the ratio of p.e. in data to isotropic simulation) are shown for the rest of the fibres in Figs. 6.5a - 6.9c.

In these polar plots, rope shadows are visible. These are caused by hold-up or hold-down ropes which the beam shines on to, casting a shadow onto the PMTs on the far side of the detector. These shadows depend on both the fibre and rope position. However, it is evident that there is a mismatch in the location of the rope shadows in some of the beam profiles. This is particularly evident in FS007, FS207, FS025, FS225 and FS193. This shows up as a bluer line of shadowed PMTs (where the rope shadow is in data) running parallel alongside a yellower line of shadowed PMTs (where the rope shadow is in simulation). This could be caused by the ropes in the detector being located in a sightly different position in the simulation or by a mismodelling of a fibre position. However, it is clear that the location of multiple SMELLIE plates would have to be wrong in the simulation, as ropes seen in the profiles of multiple different fibres across multiple different nodes are mismodelled in the profiles. This is an indication that, as this behaviour is seen all over the detector, the ropes locations may be the cause.

Visible rope shadows in the beam profiles are likely to be caused by ropes on the near side of the AV to the fibre. This was tested by running simulation with the recording of tracking information enabled and investigating the paths involving rope reflections. Therefore, estimates can be made of how incorrect the rope position
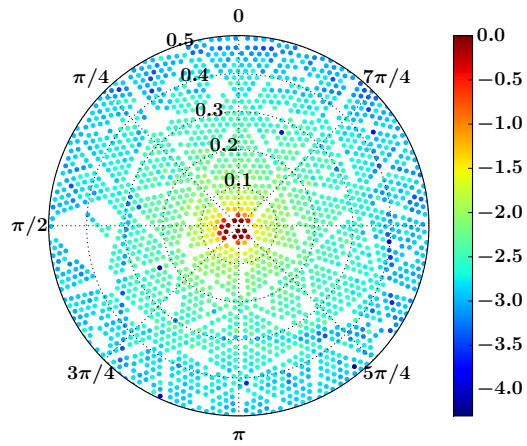
(a) Data
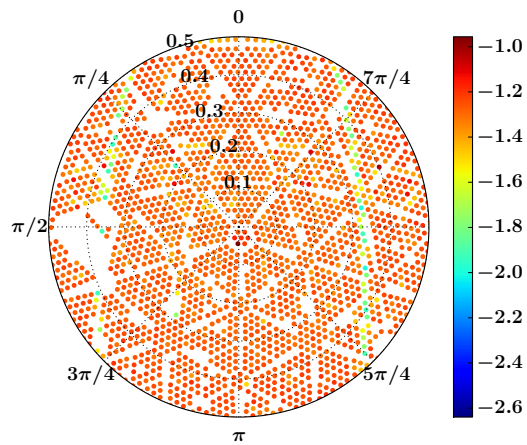
(b) Isotropic simulation
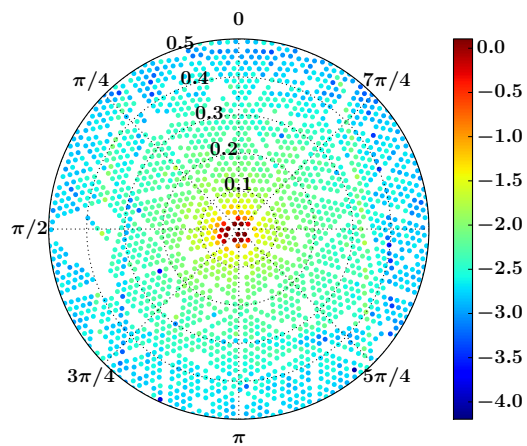
(c) Ratio

Figure 6.4: Polar plots showing the stages of creating a beam profile for FS007. (a) shows the detector data, (b) the isotropic simulation and (c) the ratio of data to simulation. The coordinate system used is described in Section 5.1. Each coloured circle represents a PMT and the colour scale, for data and simulation, shows $\log_{10}$ of the normalised number of p.e. For the ratio, it shows $\log_{10}$ of the ratio of p.e. in data to simulation.
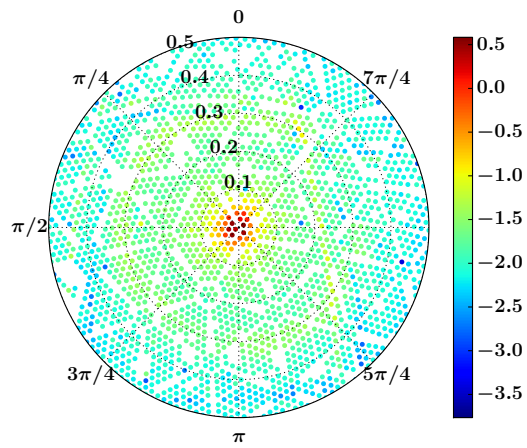
(a) FS107
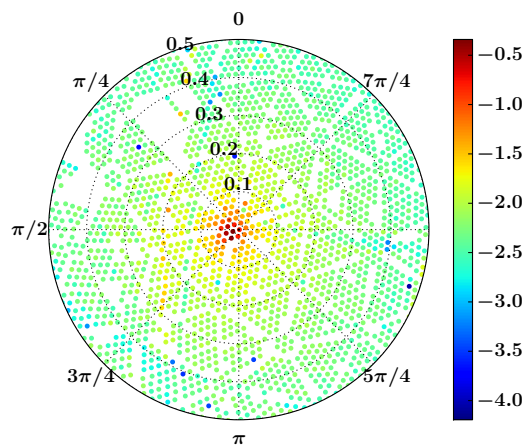
(b) FS207

Figure 6.5: Polar plots showing the ratio of p.e. per PMT in data to isotropic simulation for FS107 and FS207. The coordinate system used is described in Section 5.1. Each coloured circle represents a PMT and the colour scale shows $\log_{10}$ of the ratio of p.e. in data to simulation.
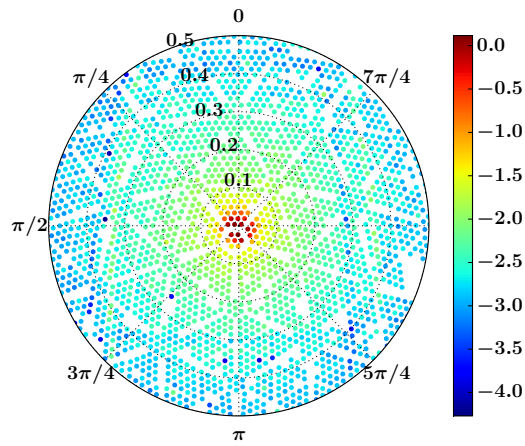
(a) FS025
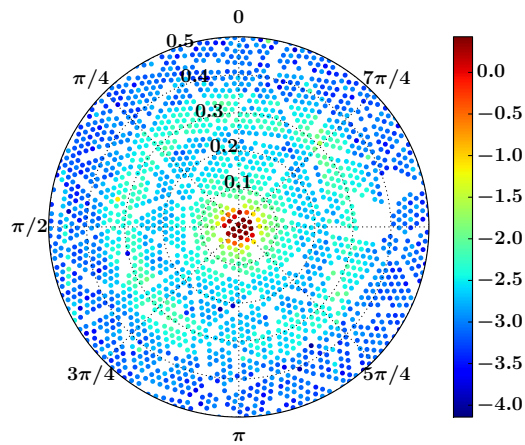
(b) FS125
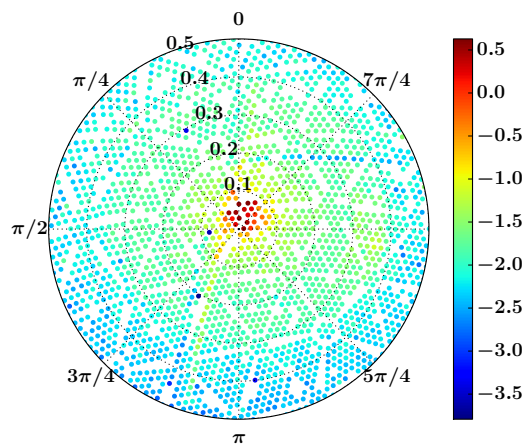
(c) FS225

Figure 6.6: Polar plots showing the ratio of p.e. per PMT in data to isotropic simulation for the fibres on the 25 node. The coordinate system used is described in Section 5.1. Each coloured circle represents a PMT and the colour scale shows $\log_{10}$ of the ratio of p.e. in data to simulation.
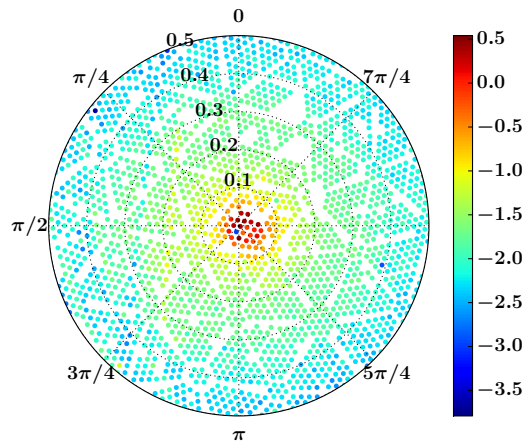
(a) FS037
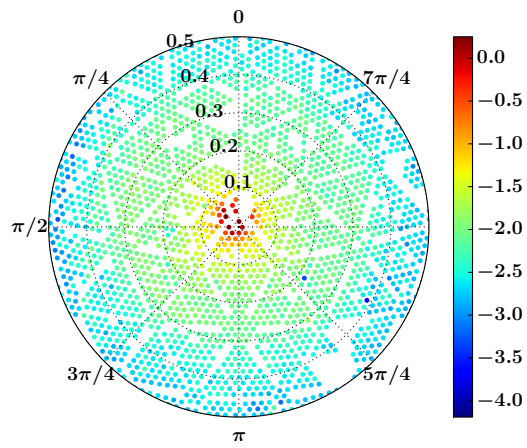
(b) FS137
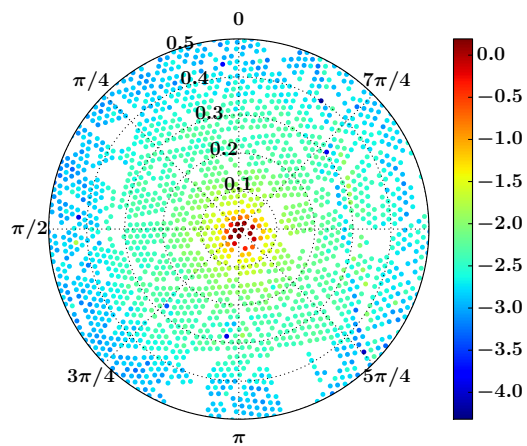
(c) FS237

Figure 6.7: Polar plots showing the ratio of p.e. per PMT in data to isotropic simulation for the fibres on the 37 node. The coordinate system used is described in Section 5.1. Each coloured circle represents a PMT and the colour scale shows $\log_{10}$ of the ratio of p.e. in data to simulation.
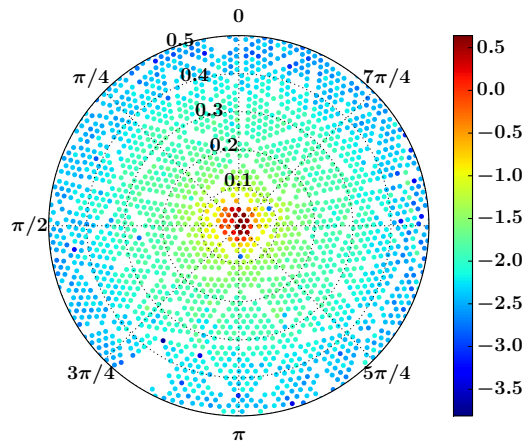
(a) FS055
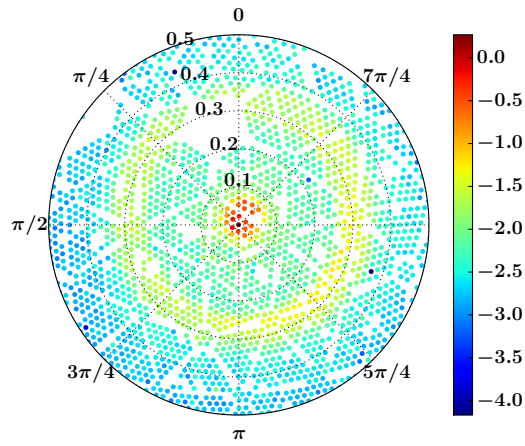
(b) FS155
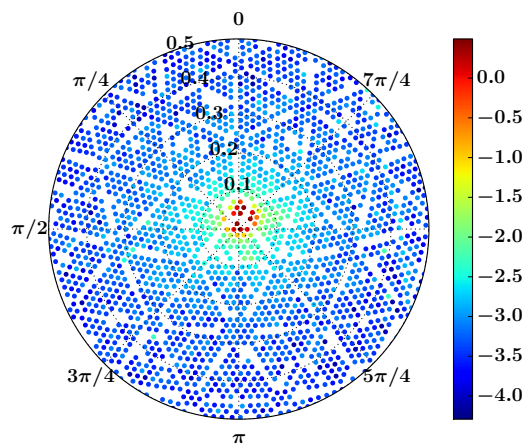
(c) FS255

Figure 6.8: Polar plots showing the ratio of p.e. per PMT in data to isotropic simulation for the fibres on the 55 node. The coordinate system used is described in Section 5.1. Each coloured circle represents a PMT and the colour scale shows $\log_{10}$ of the ratio of p.e. in data to simulation.
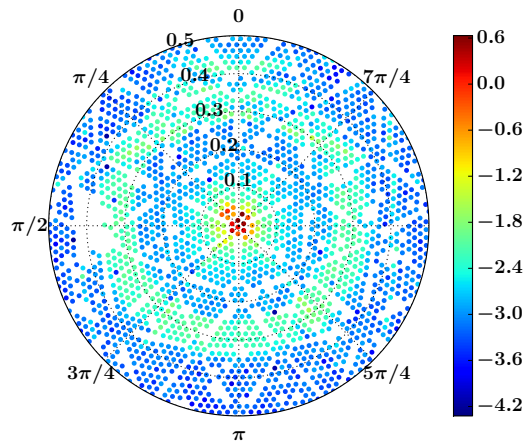
(a) FS093
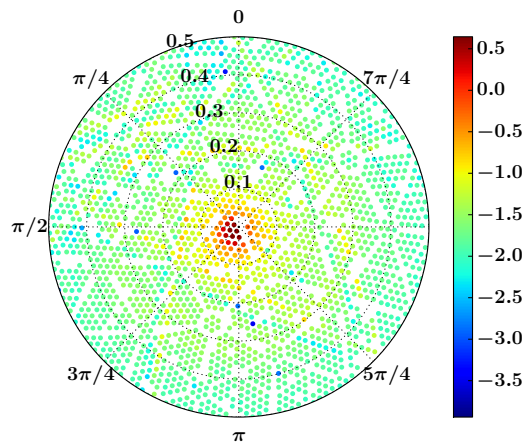
(b) FS193
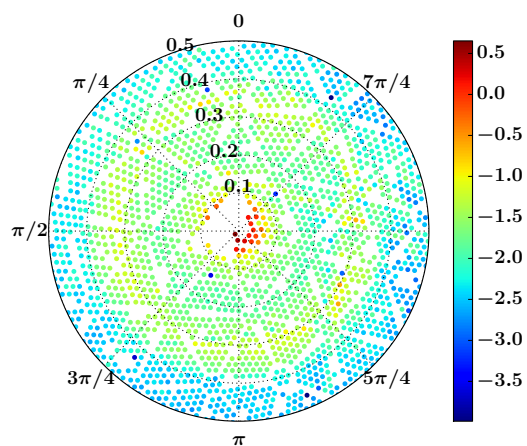
(c) FS293

Figure 6.9: Polar plots showing the ratio of p.e. per PMT in data to isotropic simulation for the fibres on the 21 node. The coordinate system used is described in Section 5.1. Each coloured circle represents a PMT and the colour scale shows $\log_{10}$ of the ratio of p.e. in data to simulation.

would have to be in order to observe the mismatch seen in the beam profiles. For the purpose of these estimates, some assumptions are made. Suppose a 0° fibre has a rope directly in front of it. A typical distance between the mismatched rope shadows is taken as approximately 4 degrees (chosen by examining the beam profiles). In this situation, the rope would be incorrectly positioned by approximately 18 cm, which, to put it in perspective, is about 1/200th of the circumference (at the equator) of the AV. It is unlikely for a fibre position to be this wrong - this would involve the SMELLIE plate being located on a different side of the hex cell than the installation team reported. It is also worth remembering that considering the case of a rope directly in front of a 0° fibre will result in the greatest displacement in rope position. In reality, the displacement will be smaller than 18 cm.

Therefore, for this work, it is assumed that the fibre positions in simulation are correct but that the rope positions in simulation do not match reality. This is not possible to correct using SMELLIE data as SMELLIE beams do not illuminate the entire detector and, consequently, a full mapping of the true rope positions is not possible. Using TELLIE for this task would be ideal and will be pursued in the future. During this work, the mismatch between rope shadow locations in simulation and data is treated as a systematic error to the scattering analysis (evaluated in Section 7.4.3).

Once the 2D profiles and generator were implemented into RAT, tests were run by simulating 10000 shots of the laser with a realistic photon number and then comparing the NHits as a function of α angle for the simulation and the matching data subrun (as done for the 1D profiles in Section 6.2.1). These comparisons are shown in Figs. 6.10 - 6.12.

It is clear that these simulations match the α distribution of data much better than the equivalent simulations created using the 1D profiles shown in Figs. 6.1 - 6.3. The largest χ^2_{DOF} of the three simulations shown here is 6.94, a reduction of two orders of magnitude from the best of the 1D simulations shown. There do remain angular sections in which there are statistically significant differences between data and simulation. In particular, simulation and data show systematic differences on the falling edge of the beam, at around 0.1 rads. In Section 7.4.2, the effect that this mismatch has on the result of the scattering analysis is investigated. It is found that this level of mismatch causes no discernible systematic error on the scattering length measured and, therefore, the differences were not investigated further.

It has previously been mentioned that only PMTs within a certain α angle of the beamspot centre are considered part of the beam profile. In addition to this, a time
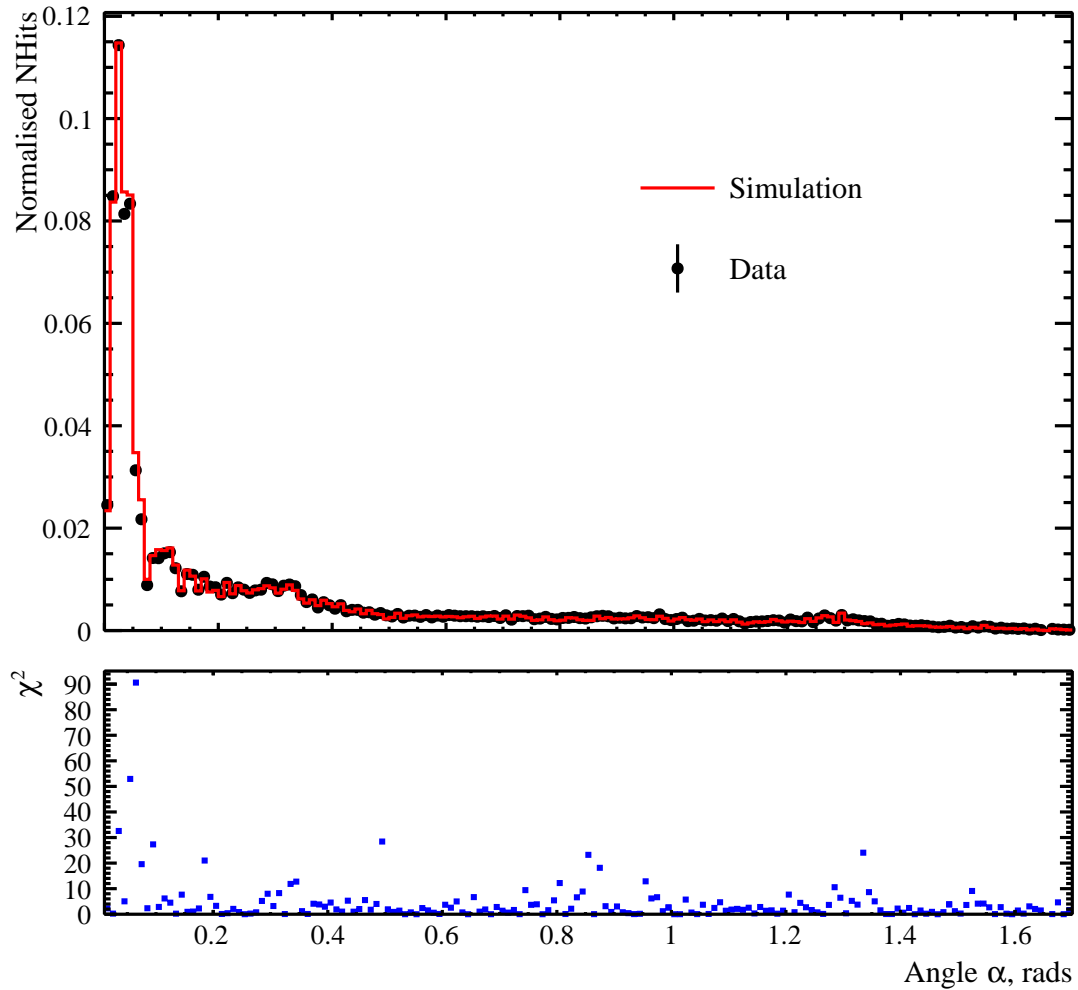
Figure 6.10: Simulation created using the 2D beam profile of FS055 at 495 nm is shown in red, compared to the matching data subrun in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 4.51.
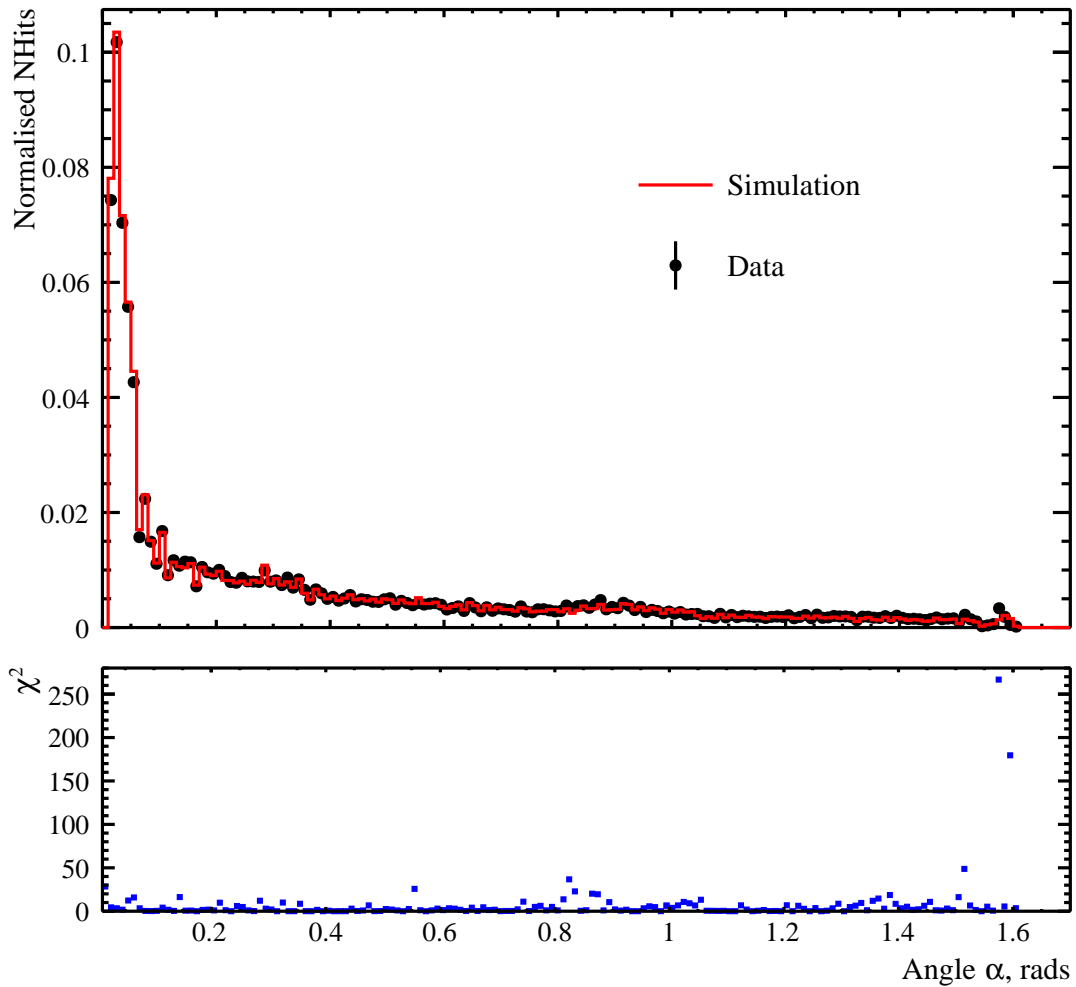
Figure 6.11: Simulation created using the 2D beam profile of FS007 at 495 nm is shown in red, compared to the matching data subrun in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 6.81.
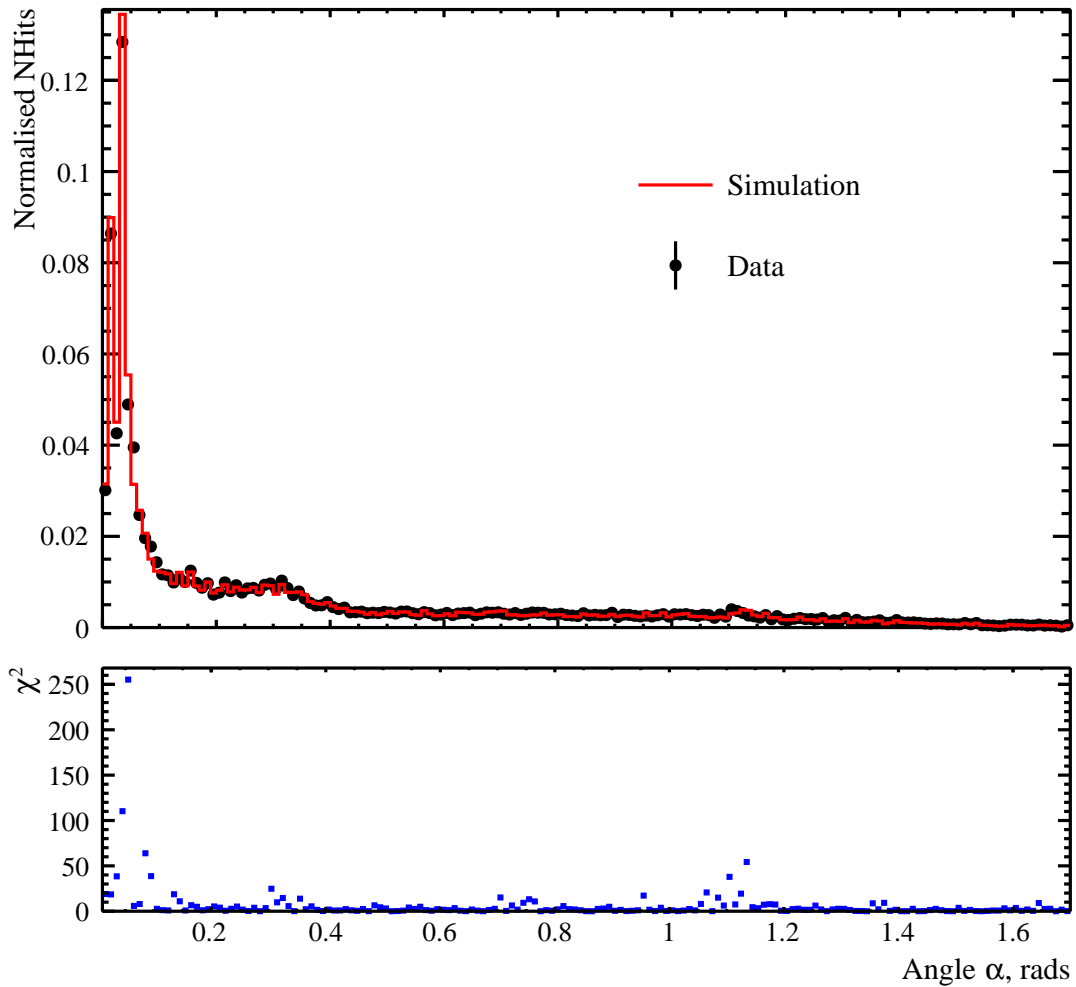
Figure 6.12: Simulation created using the 2D beam profile of FS237 at 495 nm is shown in red, compared to the matching data subrun in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 6.94.

residual cut is also used when creating the profiles such that only hits within a certain time residual window of the prompt peak are considered. (The prompt peak of the time residuals is fit with a Gaussian and the window is applied with respect to the mean of the fit.) While the angular cut clearly has to be applied to both data and the isotropic simulation, the timing cut is only applied to data. The aim of this cut is to isolate direct hits, as these are the photons that can be reconstructed into the angular distribution of photons leaving the fibre. In the isotropic simulation, however, the aim is to understand detector effects and therefore the timing cut is not necessary.

Originally, an attempt was made to decide the values for both of these cuts using a 2D grid search, in which the two dimensions were the half-width of the time residual window and the polar angle above which to exclude PMTs from the beam profile. The search was conducted by minimising the χ^2_{DOF} between data and a realistic simulation created using the beam profile in question.

However, it quickly became apparent that due to the already discussed mismatch between data and simulation at small angles, the larger the angle cut, the lower the χ^2_{DOF}, even to the point of physical impossibility. Consequently, an angular cut was chosen by eye - simply by examining the match between data and simulation at higher angles. This resulted in an angular cut of 60°.

From this grid search, it also became apparent that, provided the time residual window includes the majority of the prompt peak, the χ^2_{DOF} does not have a strong dependence on the value of that cut. Therefore, a value of 8 ns for the half-width of the window is chosen. This should ensure that at least 95% of the prompt peak will be within the window.

The profiles shown in Figs. 6.10 - 6.12 were created with these cuts and, as discussed, any mismatch between the simulation and data is negligible in terms of the scattering length analysis presented in this work.

In summary, it has been shown that the existing 1D profiles for the SMELLIE fibres were not suitable for conducting a scattering analysis. Therefore, a new 2D generator has been written purely for the purpose of simulating SMELLIE beams and 2D beam profiles have been created from detector data to use in combination with this generator. As has been demonstrated, this combination is used to create the most accurate simulation of SMELLIE beams to date. The 2D beam profiles created here are used for the remainder of this work.

Chapter 7

SMELLIE Analysis

7.1 Scattering Length Analysis

SMELLIE, as previously discussed, is designed to measure the scattering characteristics of the detector media. The most crucial measurement SMELLIE can make is the scattering length of the medium inside of the AV. This is the location of the fiducial volume for all of the SNO+ physics analyses and the scattering length is a crucial input to the optical model.

The original scattering length analysis was designed by K. Majumdar and S. Langrock and tested using simulated data sets, as the detector was not operational at this point [53, 104]. Measuring the scattering length takes the form of measuring the scaling factor (a divisive factor applied to the Rayleigh scattering lengths, defined in Chapter 3). This means that a direct comparison between data and Monte Carlo is drawn and the scattering length in data is measured by adjusting the default scaling factor in RAT until it best represents the data.

In this work, the analysis has been extended to be suitable for detector data and implemented across a range of fibres and wavelengths, in order to calculate the scaling factor of the internal water. The analysis will first be presented for a single fibre/wavelength combination, before moving on to the application across all the fibres and a range of wavelengths. The analysis will also be used to determine information regarding other optical properties of the detector, such as the scattering length of the acrylic. Finally, the various systematic errors that influence the scattering length measurement are investigated.

7.1.1 A Single Fibre at a Single Wavelength

As discussed, firstly, the scattering length analysis will be explained using a single fibre/wavelength combination. The fibre used for this is FS007, which is a 0° fibre (with no ropes through the centre of the beamspot) and 495 nm light from the superK laser. This has been chosen as an example of the analysis in a simple case. The data used is subrun 139 of run 114018. This is a medium shot energy (and hence multi-hit correctable) subrun and is taken from the same run that the beam profiles were made from, meaning that the detector conditions have not changed. No unusual behaviour is seen during this subrun: the CAEN traces do not show any sign of laser instability, no HV trips occurred, no FIFO overflows and no activity on deck. 495 nm is chosen as this wavelength was used to create the beam profiles and the availability of a PQ laser at the same wavelength provides the opportunity to run checks with a hardware configuration.

The optics used for this analysis are the original RAT optics described in Section 3.2.1. When this work was started, this set of optical constants was the only available set and hence were used by default.

The first step of the analysis is to calculate the number of photons fired into the detector from the fibre, in order to be able to generate matching simulations. This is done by calculating the number of p.e. in the chosen data subrun. Run-level simulations are performed with varying numbers of photons (500 - 1500 photons in 5 steps). This range of photon numbers was chosen from previous experience calculating p.e. during the building of the tesseract. Using run-level simulations means that the detector state is simulated as accurately as possible, including noise hits, on a PMT by PMT basis. Therefore, there is no need to include the corrections for screamers and other noise hits that were used in the tesseract data cleaning.

The mean number of p.e. per event summed over the entire detector in each of these simulations is calculated. This is done by calculating the occupancy of each PMT during the subrun, multi-hit correcting the occupancies, then summing the p.e. across all the PMTs.

The number of p.e. per event is fitted as a linear function of the mean number of photons fired into the detector in simulation, as shown in Fig. 7.1. The fit is then used along with the number of p.e. observed in the detector in the data to calculate the number of photons fired into the detector in data. In this case, the number of photons is calculated to be 1271 ± 4 (stat.).

The statistical errors on each point in Fig. 7.1 are simply Poisson errors - the square root of the mean number of p.e. per shot divided by the square root of the number of
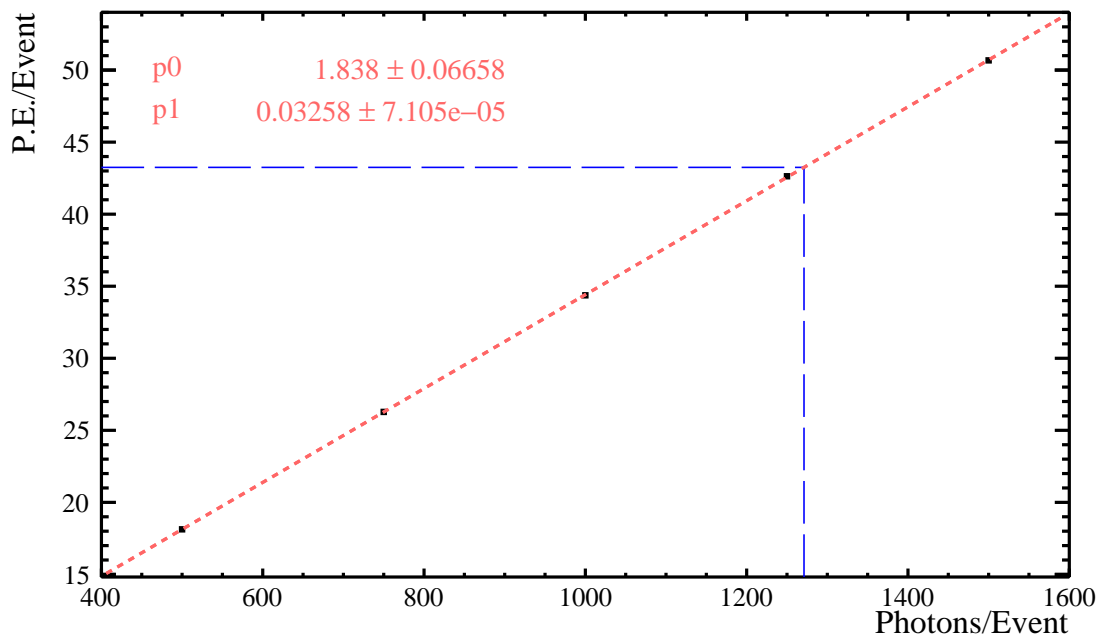
Figure 7.1: Number of p.e. detected per simulated event as a function of mean number of photons per pulse fired into the detector in simulation. The fit (in red) is a straight line of the form $y = \text{p0} + x \times \text{p1}$. The results of the fit for p0 and p1 are shown in the top left. The p.e. per event in the matching data subrun is calculated and the fit is used to translate this into the number of photons fired into the detector in data (blue lines).

shots. This is an estimate of the error as the occupancy is an average measurement. The systematic error arising from non-Poisson variation in photon number per shot will be explored later in this chapter, in Section 7.4.6. The statistical error on the photon number in data is the combination of the Poisson error on the number of p.e. per event in the data subrun and the error from the fit.

As already discussed, each data subrun consists of 10000 shots in which the laser intensity is tuned to give a maximum occupancy of approximately 0.8. The simulations used to calculate the photon number consist of 1000 shots. This reduction in statistics was introduced purely to minimise the computation time required for the simulations. Testing was done to ensure that this did not introduce an unexpected systematic error on the calculated photon number in data. It was shown that the resultant change was completely contained within the statistical errors.

Following this, run-level simulations are run with the number of photons per shot as calculated above (in this case 1271). Instead of varying the photon number fired, the Rayleigh scattering scaling factor is varied between 0.5 – 1.5 in five regular steps. This range was chosen as it represents a $\pm 50\%$ error in the RAT scaling factor. This is a believable range - the RAT scaling factor is not expected to be orders of magnitude wrong.

The simulation with scaling factor equal to one has the option to save the photon tracking information turned on. This means that each photon emitted can be traced from its emission to its demise, whether that is entering a PMT, absorption in one of the detector media or exit from the detector. In practice, due to the number of possible combinations of interactions, only certain categories of photon interactions are of interest in this analysis. Crucially, only photons which enter a PMT are considered. These are divided up into the following general categories (where optical interactions mean scattering, reflection and absorption; refraction through the AV is assumed but not considered as a separate optical interaction):

- **Internal Water Scatter** - Photons which scatter once in the internal water and undergo no other optical interactions.

- **Direct** - Photons which undergo no optical interactions before entering a PMT.

- **Scatter in AV** - Photons which are scattered once in the acrylic but undergo no other optical interactions.

- **External Water Scatter** - Photons which undergo a single scatter in the external water but no other optical interactions.

- **AV Reflection** - Photons which are reflected once from one of the interfaces between the water and the AV and undergo no other optical interactions.

- **PMT Reflection** - Photons which are reflected once from either a PMT concentrator or PMT glass.

- **Multiple Effects** - Photons which undergo more than one optical interaction, including more than one scatter, multiple reflections or a combination of both.

- **Noise** - Hits which do not correspond to a photon belonging to any of the above categories, caused by detector noise.

Each detector hit is matched to an optical category. When classifying each photon to a category, the ID of the PMT the photon entered is recorded. Separately, the PMT hits are looped through and their time residuals for each hit and α angles calculated. This is done event by event. Therefore, if, in one event, a PMT only has one photon enter it, it is trivial to match the photon to the PMT hit. However, SMELLIE events are not in a single photon regime - hence the necessity for the multi-hit correction. Therefore, if more than one photon enters a PMT during a single event, the last photon stored in the data structure which enters that PMT is used to classify the optical category of the PMT hit.

Fig. 7.2 shows the photons which entered PMTs from such a simulation. The photons are split into the above categories. Each photon is represented by a dot, coloured by the category it belongs to (yellow for internal water scatter, red for direct, black for noise, etc.). The photons are plotted in 2D - on the x-axis is the α angle of the PMT (see Chapter 5.1) and the y-axis is the time residual (as discussed in Chapter 2) of each hit.

This method does result in some misclassified photons. As discussed, it is possible for multiple photons to enter a single PMT in any given event. However, a PMT can only record one hit time per event. As stated above, only one hit time is associated with all of those photons. Therefore, this leads to photons in Fig. 7.2 in the wrong vertical position. This is easily visible in cases when photons of particular optical categories are in positions on the the y-axis which are physically impossible. One example is direct photons at times when only noise hits are possible (it is too early in time residual for any photon exiting the laser to have reached a PMT). However, this misclassification is negligible to the analysis and therefore a more sophisticated treatment is not necessary. It has been calculated that less than one PMT per event
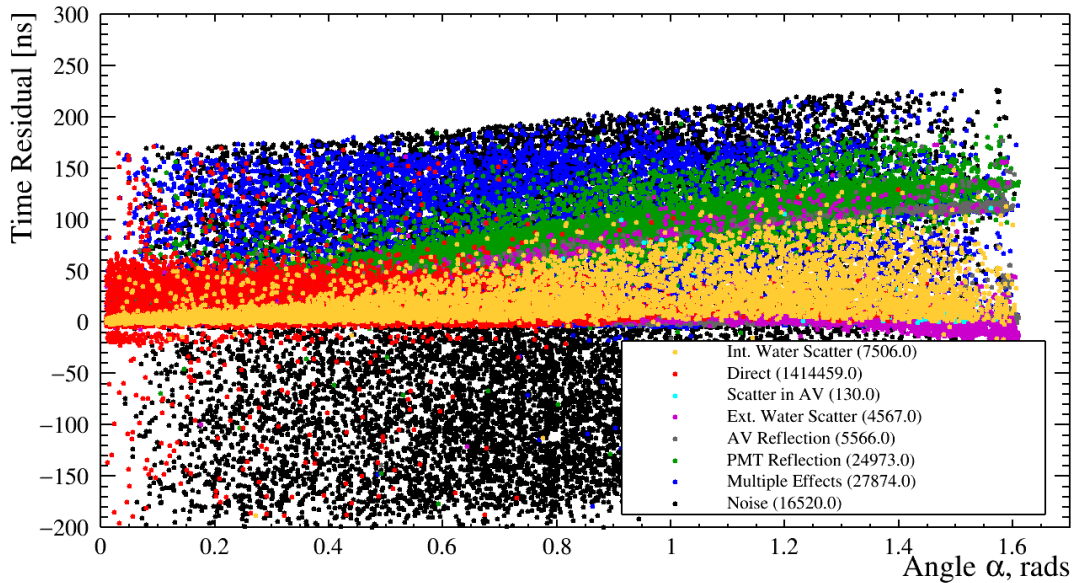
Figure 7.2: The time residual versus α angle of each PMT hit in a simulated SMELLIE subrun (FS007, 495 nm with the superK). Tracking information is used to classify each hit into a category based on the optical interactions the photon has undergone between emission and entering a PMT. For example, red dots indicate hits caused by photons which underwent no optical interactions. The legend denotes all of the categories and the number in brackets is how many photons are in that category.

at angles in the selection region (which will be discussed below) has more than one photon enter it.

While it is useful to see how these populations of hits relate to each other in this time versus angle space, it is difficult to fully appreciate the distribution for each population when they are overlaid. Therefore, in Figs. 7.3 - 7.10 are the time residual vs. angle plots for each category. The number in the legend, next to the optical category, refers to the number of photons in the category over all events in the subrun.

From these plots, it is apparent that there are strong patterns in the distributions. These are caused by the detector geometry. The two lines shown in the distributions (such as in Figs. 7.6 and 7.7) are caused by the two regions of AV and external water, one close to the fibre and one on the far side of the AV, near the beamspot on the PSUP. For example, take the photons scattered once in the external water (Fig. 7.6). Photons, once scattered, are able to hit any PMT, hence there being hits at all angles. However, small angle scatters dominate Rayleigh scattering. These will have a time residual that is close to a direct hit as the path has not diverged far from that of a direct hit. These hits are the corner of the triangle in Fig. 7.6 near a time residual of
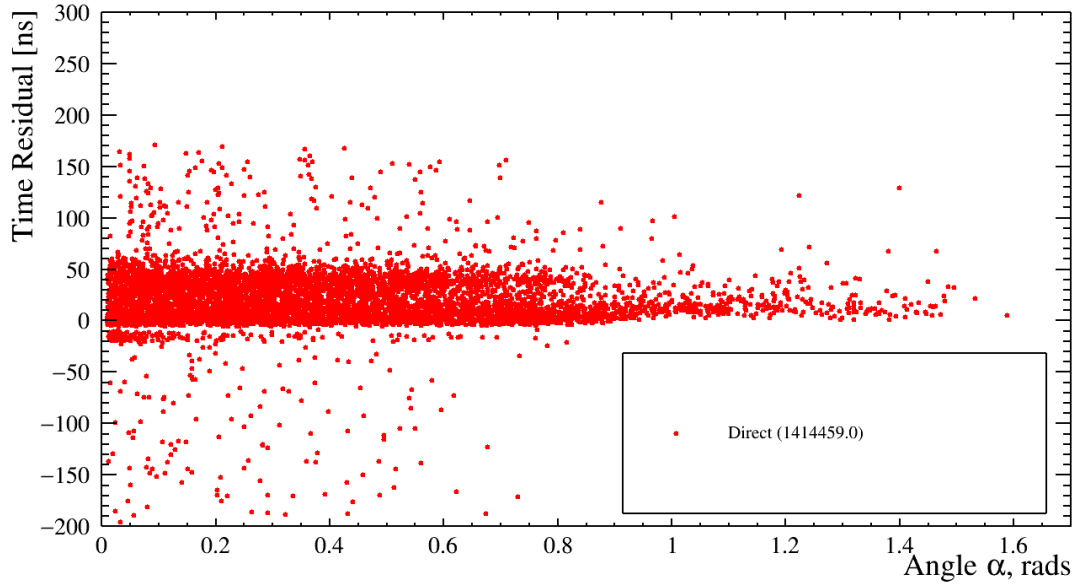
Figure 7.3: Time residual versus α angle for photons having reached PMTs directly in the simulated SMELLIE subrun shown in Fig. 7.2.
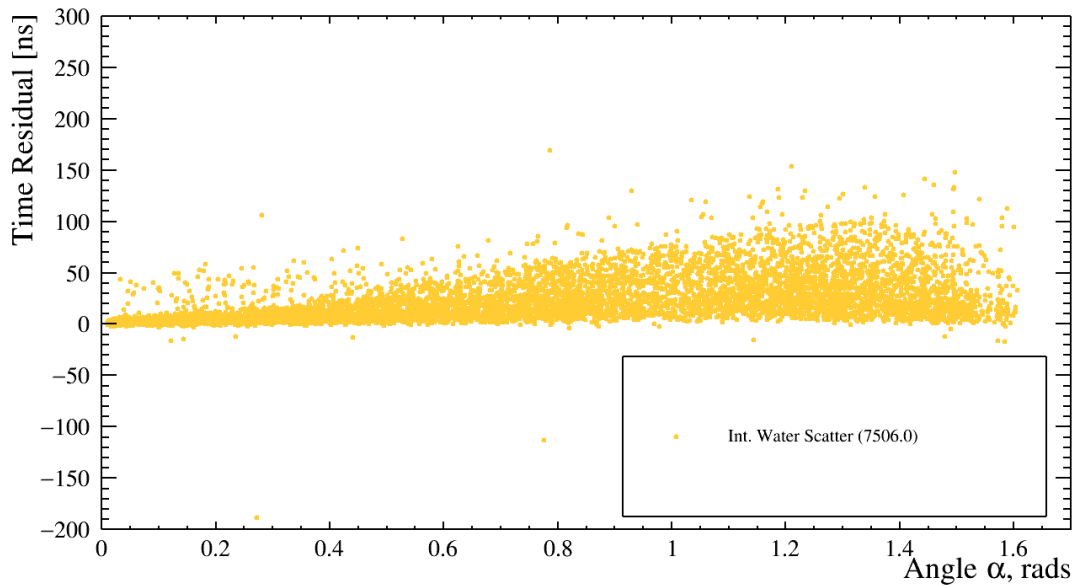
Figure 7.4: Time residual versus α angle for photons singly scattered in the internal water in the simulated SMELLIE subrun shown in Fig. 7.2.
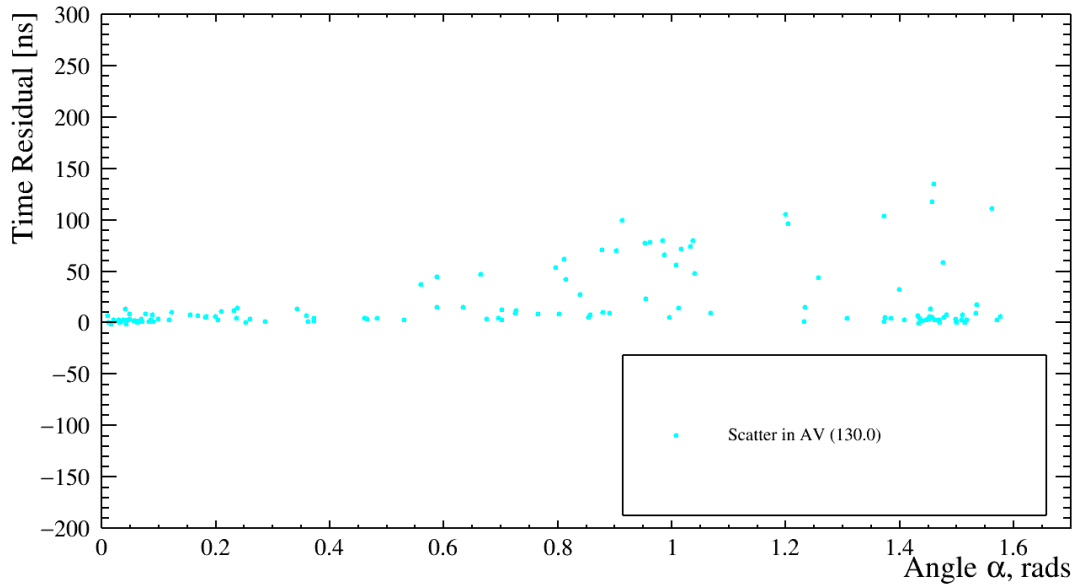
Figure 7.5: Time residual versus α angle for photons singly scattered in the acrylic in the simulated SMELLIE subrun shown in Fig. 7.2.
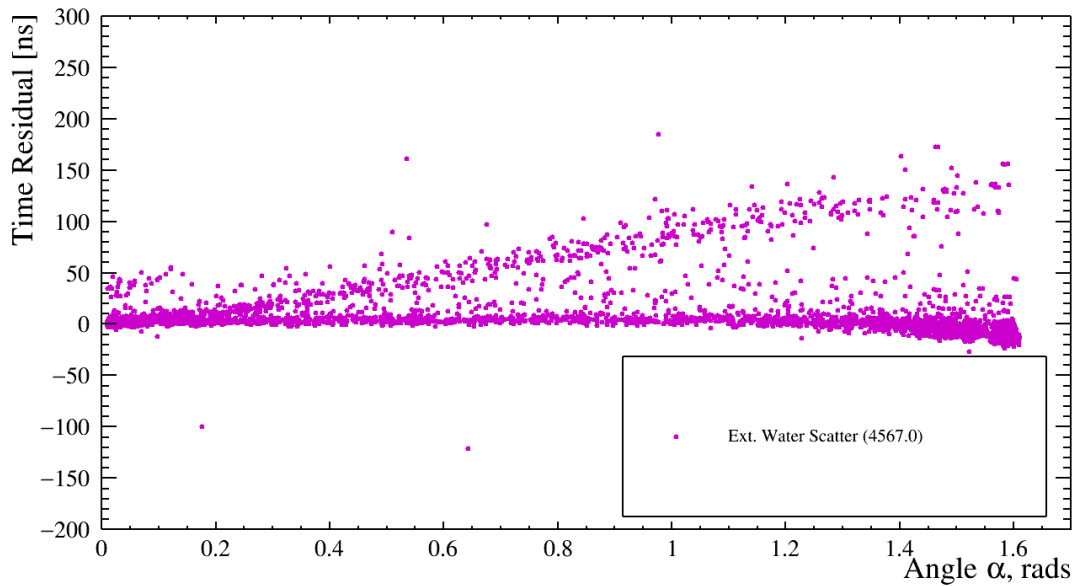
Figure 7.6: Time residual versus α angle for photons singly scattered in the external water in the simulated SMELLIE subrun shown in Fig. 7.2.
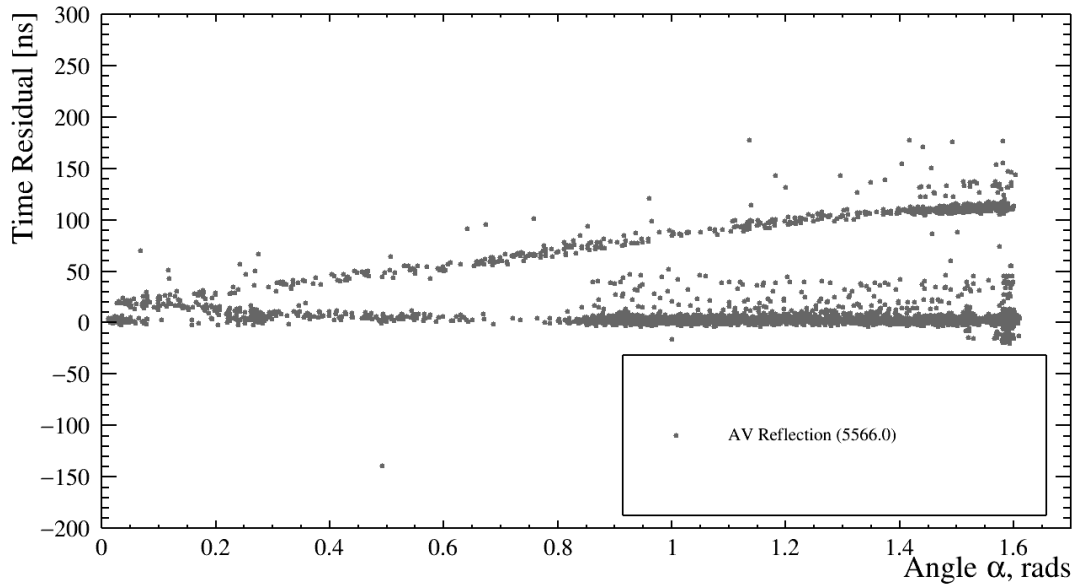
Figure 7.7: Time residual versus α angle for photons reflected by an AV/water interface in the simulated SMELLIE subrun shown in Fig. 7.2.
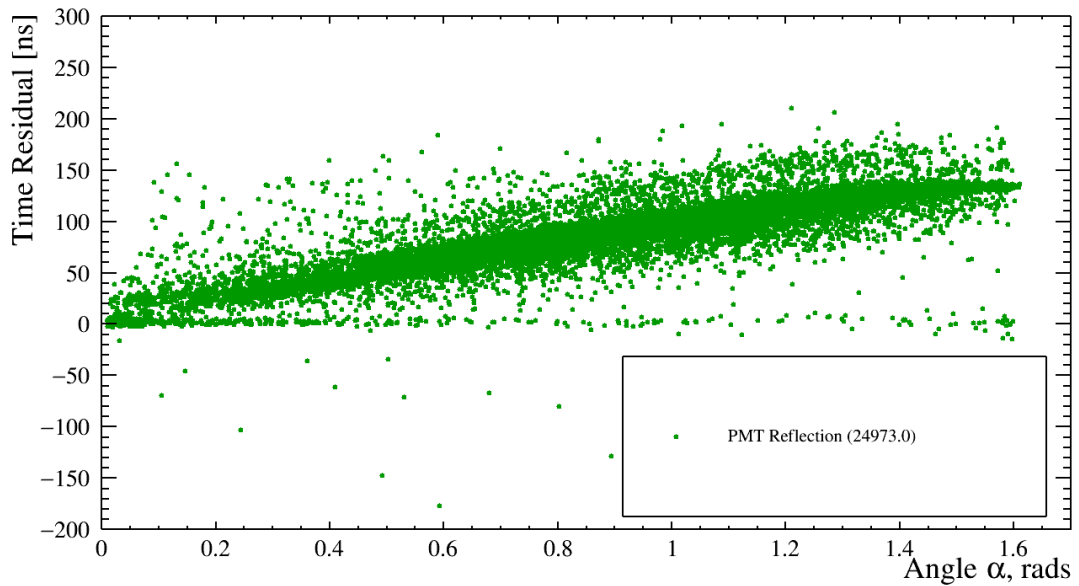
Figure 7.8: Time residual versus α angle for photons reflected from a PMT in the simulated SMEL-LIE subrun shown in Fig. 7.2.
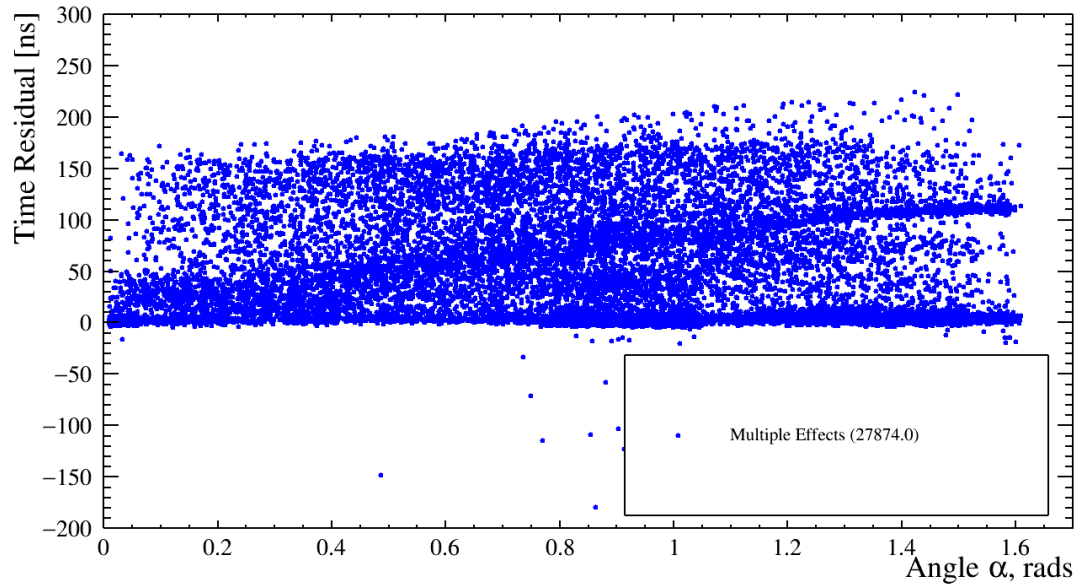
Figure 7.9: Time residual versus α angle for photons having undergone more than one optical effect in the simulated SMELLIE subrun shown in Fig. 7.2.
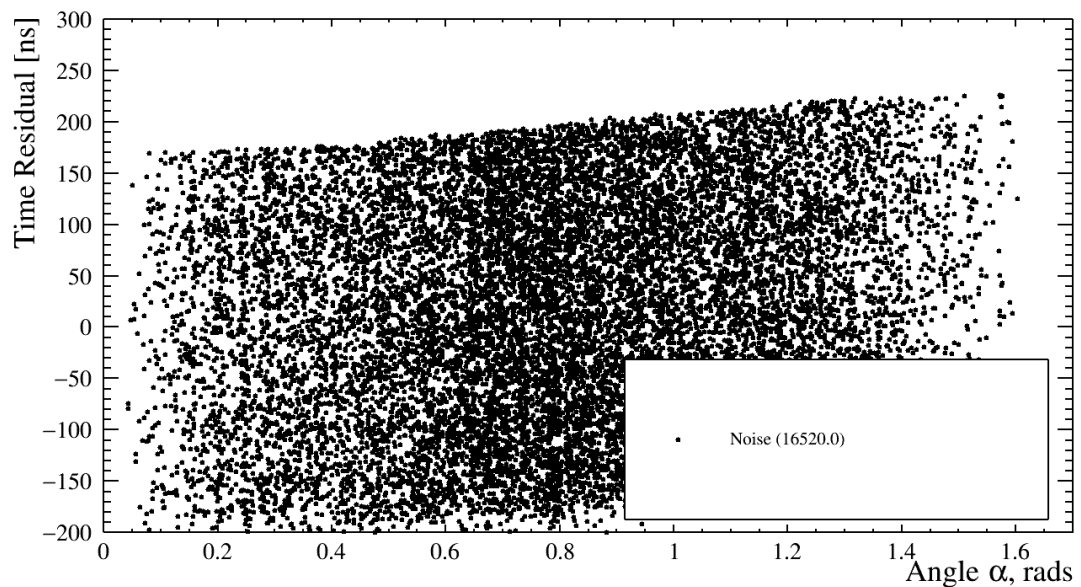
Figure 7.10: Time residual versus α angle for hits caused by detector noise in the simulated SMELLIE subrun shown in Fig. 7.2.

0 ns and an angle of 0 rads. Equally, high angle scatters which occur in the external water before the photon has crossed the AV look (in terms of hit location) like a reflection from the AV. Due to this, the `LightPathCalculator` assumes these hits are caused by AV reflected photons and therefore the time residual for these hits become marginally negative. (The time residuals for these particular hits are too early to be anything other than scattering in the external water.) These are the hits in the bottom line of hits in Fig. 7.6 on the right side of the plot. However, high angle scatters from photons which have already crossed the AV are also modelled as being caused by photons undergoing AV reflection. This causes a large time residual (on the order of 100 ns). Such logic, extended across all angles, explains the two lined triangular structure of the distribution in this 2D space.

The signal in this analysis is considered to be photons in the category of a single scatter in the internal water. Everything else is a background to this. A selection of hits is chosen with the greatest $\frac{s}{\sqrt{s+b}}$, where s is the number of signal hits selected and b background hits. Due to the geometric nature of the distributions in this 2D space, this region of hits (henceforth known as the selection region) is selected using two straight lines, forming two sides of a triangle.

Using two lines (and the line of $\alpha = 1.7$ rads) to select a region relies upon the ideal region being a single, triangular area in the 2D space of time residual vs. angle. Initially, this seems a reasonable assumption due to the distribution of singly scattered hits in this space, as shown in Fig. 7.4. However, the distribution of other categories of hits affects whether these singly scattered hits have a favourable $\frac{s}{\sqrt{s+b}}$. The $\frac{s}{\sqrt{s+b}}$ is calculated per bin, as shown in Fig. 7.11. The bins used are 1 ns by 0.01 rads. Small bins are used as it is not desirable to have multiple photons falling into the same bin in the same event. Therefore, the angle divisions have been chosen to match the size of the PMTs. Clearly, the angular size of a PMT as viewed from a fibre depends on the location of the PMT relative to the fibre. A PMT directly opposite the fibre is the smallest and a PMT next to the fibre is the largest. Therefore, the bins roughly match the angular size of the smallest PMT.

The gradients and intercepts of the two lines that form the selection region are chosen using a 4D grid search, in which the four dimensions are the two intercepts and the two gradients. The aim of the search is to find the combination of these dimensions that results in the largest $\frac{s}{\sqrt{s+b}}$ contained within the selection region. The grid search parameters are the boundaries on each of the dimensions and the number of steps in each dimension (the combination of which determines the step
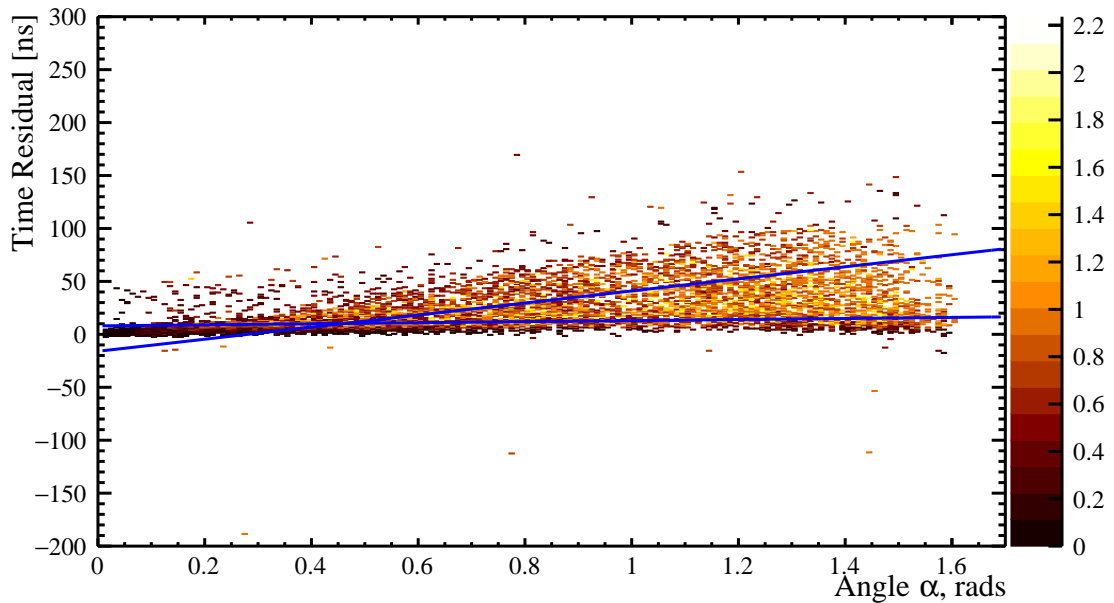
Figure 7.11: Time residual versus α angle plot with each bin coloured by $\frac{s}{\sqrt{s+b}}$, where s represents the signal - singly scattered photons - and b represents any other photon, for the simulated SMELLIE subrun shown in Fig. 7.2. The colour scale is shown to the right of the plot. If there is no signal in a bin, the bin is white. The two blue lines (and the right-hand side of the plot) show the result of the grid search to find the selection region.

size). Investigations were performed into the effect changing each of these parameters had on the resultant region.

Using Fig. 7.11, the optimal values of the intercepts and gradients can be estimated. For example, the lower line is expected to have a gradient and intercept relatively close to zero. Therefore, this intuition was used to set the bounds of the dimensions initially. The bounds were then varied to test how much the results changed. The effect of the resultant levels of change (albeit for a slightly different case) is explored in Section 7.4.5.

In addition, the 4D space was explored through 1D and 2D projections. It is difficult to fully understand the topology of the space through projections, but they were used to ensure that the boundaries enclosed the optimal values of each dimension and to check for signs of oversampling. An added complication is that the four dimensions are highly correlated. This is not surprising - for example, if the intercept of one of the lines is decreased, the optimal value for the gradient of that line will be increased.

Fig. 7.11 also shows the results of this 4D grid search. The lines in blue show the optimal values of each dimension and the area above approximately 0.45 rads enclosed between them is the selection region.
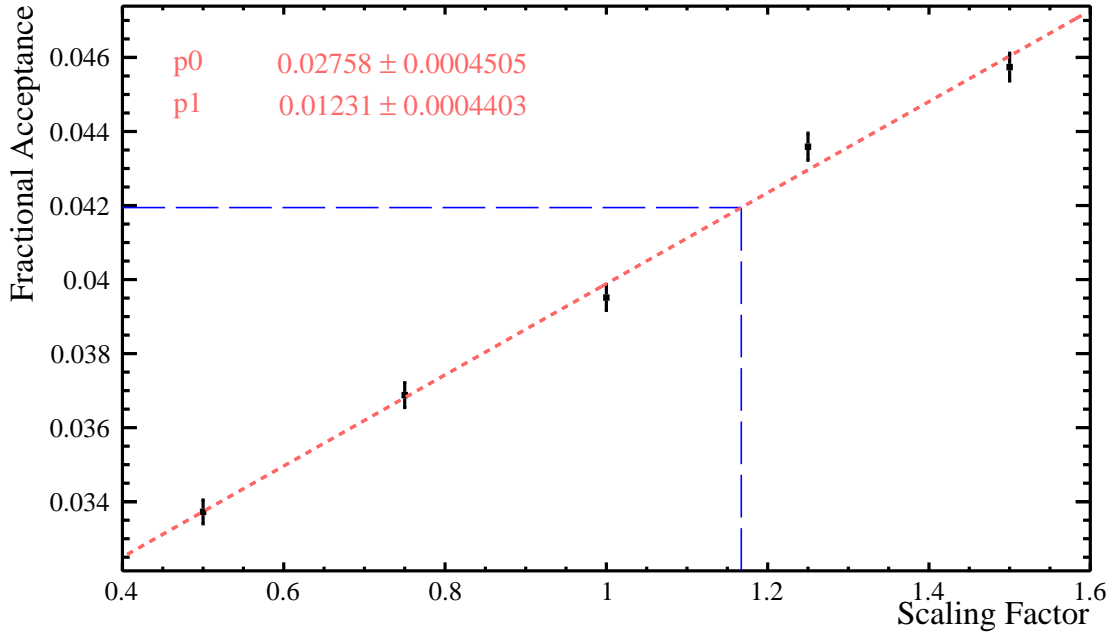
Figure 7.12: Fractional acceptance (number of hits in the selection region divided by the total number of hits in the subrun) for each simulated subrun as a function of the scaling factor used in that simulation. The fit (in red) is a straight line of the form $y = \text{p0} + x \times \text{p1}$. The results of the fit for p0 and p1 are shown in the top left. The fractional acceptance in the matching data subrun is calculated and the fit is used to translate this into the scaling factor that best matches in (show by the blue lines).

Once the optimal selection region has been found (using the simulation with a scattering scaling factor of one), it is applied to the rest of the simulations with scaled scattering lengths. For each of these, a quantity is calculated which has a strong correlation with the scaling factor. This quantity is the number of hits in the selection region divided by the total number of hits in the 2D space and will be referred to from here on as the fractional acceptance. This is shown in Fig. 7.12, where the fractional acceptance has been plotted against the scaling factor used in the simulation and has been fitted with a linear function. In this case (FS007 at 495 nm using the superK), the linear function is $y = (0.0276 \pm 0.0005) + x \times (0.0123 \pm 0.0004)$.

Next, the selection region is applied to the subrun of data and, again, the fractional acceptance is calculated. Using the linear function above, the scaling factor that corresponds to the fractional acceptance in the data can be calculated. This is shown using the blue lines in Fig. 7.12. For the subrun that has been used throughout this section, the result is a scaling factor of 1.16 ± 0.05. As before, the error on this quantity is purely statistical. This statistical error arises from the statistical error on the number of hits in the selection region, the statistical error on the total number

of hits in the 2D space (these determine the error bars on each point in Fig. 7.12) and the error from the fit. The errors on the numbers of hits, both in the selection region and in the whole 2D space, are Poisson errors. All these errors are combined assuming they are independent. Obviously, the two Poisson errors are not completely independent, as the hits in the selection region are contained in both quantities. However, the two numbers are large enough and different enough that treating them as independent is valid.

One concern regarding this analysis is that the selection region is found once using the simulation that has a scaling factor of one. This region is then applied to both simulations in which the scaling factor is different and data, in which it is unknown. In order to measure the resultant systematic error, simulations with tracking enabled were run using a variety of scaling factors and the selection regions were found. The level of change in the selection regions is the same as discussed in Section 7.4.5.

For the scaling factor simulations, 10000 shots are used to match the data. However, photon thinning has been applied. This is a feature of RAT in which a thinning factor can be specified in the simulation macro. The number of photons requested is then divided by this factor. To make up for fewer photons, the efficiency of the PMTs within RAT is multiplied by the same factor [110]. This is designed to minimise the computation time involved in tracking photons through the detector. However, there are limits to the photon thinning factor, as the overall efficiency of the PMT cannot be greater than one [65]. In this analysis, the thinning factor is always found with respect to 439 nm, where the PMT efficiency is at its highest. This ensures that the efficiency of the PMT never exceeds one.

It has been tested whether using photon thinning in the case of this analysis introduces any systematic error to the final result. This was done at 495 nm with FS055 by performing the analysis once using photon thinning on the simulation and once without. The difference in the two scaling factors found was smaller than the combined statistical error.

7.1.2 Application to all Fibres

The analysis so far has only been applied to a single fibre, FS007. However, using all the fibres results in a more robust analysis as the fibres have been installed such that the photon paths from different fibres involve different distances through the internal water as well as those paths passing through different regions of the internal water.

The process to calculate the number of photons to fire into the detector in the scaling factor simulations is repeated for each fibre and the results are shown in

Fibre	Photon Number
FS007	1271±4
FS107	1161±4
FS207	1158±4
FS037	1186±4
FS137	1396±4
FS237	1107±4
FS025	986±3
FS125	1104±3
FS225	1213±4
FS055	1029±3
FS155	1248±4
FS255	898±3
FS093	1276±4
FS193	3927±10
FS293	3766±12

Table 7.1: The number of photons calculated for 15 data subruns, with each subrun using a different SMELLIE fibre. These numbers are found by calculating the number of p.e. in simulations with set numbers of photons fired into the detector, fitting the results and using that fit to calculate how many photons were needed to cause the number of p.e. seen in the data subrun. The errors are statistical errors only.

Table 7.1. The errors shown in this table are statistical, as before. 13 of the subruns have photon numbers of approximately 1000, which makes sense given that the laser intensity for each subrun was adjusted to aim for the same occupancy regardless of fibre. Two of the subruns (those associated with FS193 and FS293), however, have significantly higher photon numbers. This is not completely unexpected as FS193 and FS293 have beams that are heavily shadowed by the hold-down rope net and the neck of the AV.

Initially, the plan was to apply the analysis completely independently to all the fibres, with selection regions being optimised for each fibre separately. However, it was decided that the same region should be used for every fibre.

To decide which region to use, the regions from all fibres at the same pointing angle were compared and the most conservative combination of gradients and intercepts were chosen. This involved picking the smallest values for the intercept and gradient of the upper cut and the largest values for the lower cut. This ensures that the conservative region is fully contained in every fibres' individually optimised selection region. A conservative region was used in order to minimise the error introduce by using an identical region across fibres - the loss of statistics involved in cutting
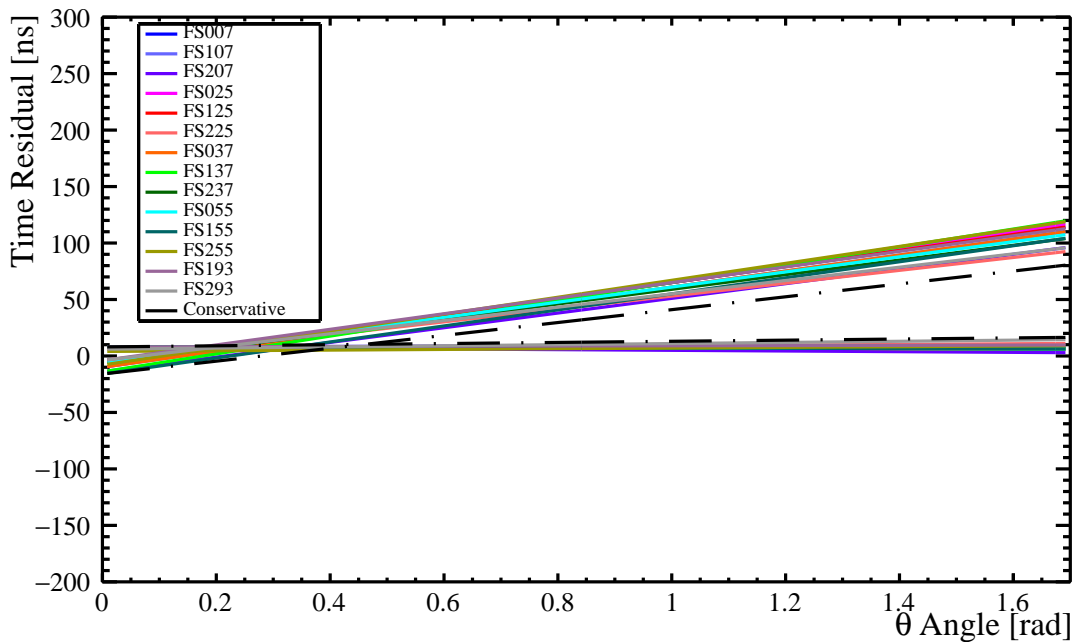
Figure 7.13: The selection regions as individually optimised for each fibre are shown in different colours. The conservative region (shown with a black dashed line) is chosen by taking the smallest intercept and gradient found for any of the upper cuts and the largest intercept and gradient found for any of the lower cuts. The scales are the same as the previous figures for comparative purposes.

out a proportion of the signal causes a smaller error than increasing the amount of background in the selection region. For example, using the region optimised for FS007 gives a percentage error of 4.3% on the scaling factor (using the data subrun associated with FS007). Using a conservative region gives a percentage error of 5.6%. However, using a region which is believably larger than the optimal region (chosen by selecting gradients and intercepts from other fibres) results in a percentage error of 6.5% on the scaling factor. This is caused by the distribution of the hits in this 2D space. Moving to a more conservative region does result in the loss of signal, but moving to a larger region begins to introduce the more dense populations of direct hits, external water scatters and AV reflections.

Each pointing angle was treated separately to be cautious. However, there did not appear to be a clear correlation between the fibre pointing angle and the selection region chosen. Therefore, the most conservative region across all fibres was chosen. The regions selected for every fibre individually and the resultant conservative region for all fibres at 495 nm using the superK are shown in Fig. 7.13.

However, one fibre is emitted from Fig. 7.13. When conducting the individual scattering analyses, the result from FS093 was far outlying, at -0.28 ± 0.1. Investigating
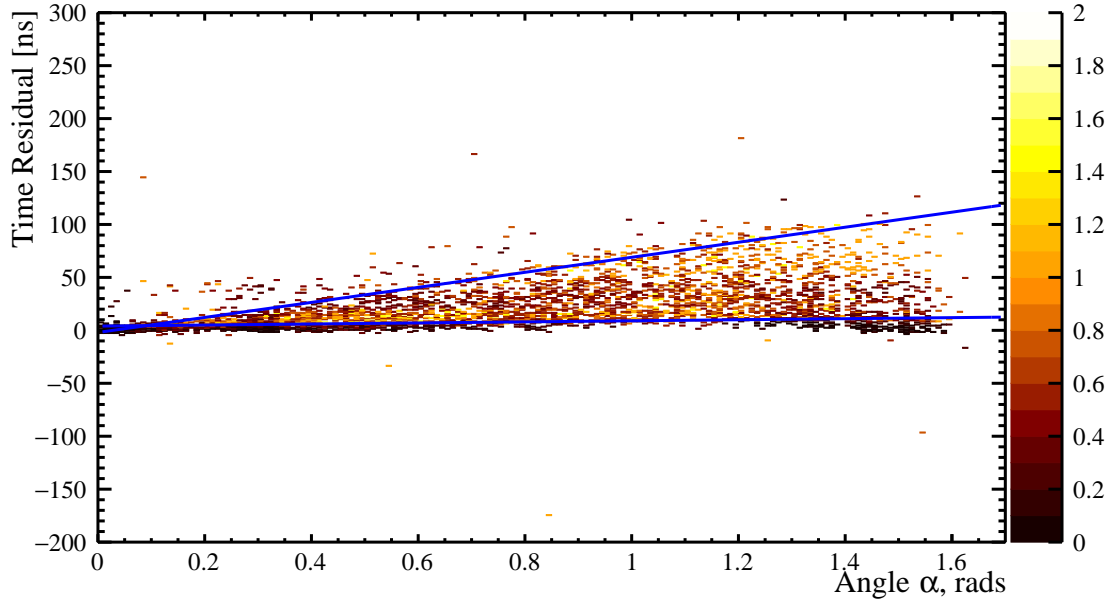
Figure 7.14: Time residual versus α angle plot showing $\frac{s}{\sqrt{s+b}}$ in each bin in a simulated SMELLIE subrun of FS093. The colour of the bin corresponds to the $\frac{s}{\sqrt{s+b}}$ in that bin and the colour scale is shown to the right of the plot. The two blue lines show the selection region chosen.

this revealed that there are extremely dominant rope reflections contaminating the analysis region for this fibre. This is not entirely unexpected. When deciding the placement of the 93 node fibres (actually installed on node 21), it was known that these fibres would be more susceptible to rope shadows than the other positions. This is due to the hold down rope net located around the neck of the detector. However, installation was continued due to the insights that near vertical paths could potentially provide on vertical variations of optical properties or precipitation.

The selection region optimised for FS093 is shown in Fig. 7.14. This is clearly not suitable - the majority of bins within the region have a $\frac{s}{\sqrt{s+b}}$ of under one and there is no longer a clear optimal region distributed as a triangular shape. However, for the purposes of this thesis, there are three other fibres with a pointing angle of 0° and 14 other fibres in total. Therefore FS093 was simply not considered from this point forward. This means that the region optimised for FS093 is not taken into account when forming the conservative region.

The systematic error introduced from altering the selection region from an individually optimised one to a conservative one is explored in Section 7.4.5. Therefore, it will not be discussed here.

The scaling factors calculated for each fibre are shown in Table 7.2 alongside their

Fibre	Scaling Factor
FS007	1.25±0.07
FS107	1.30±0.10
FS207	1.38±0.16
FS037	1.11±0.07
FS137	1.14±0.06
FS237	0.97±0.07
FS025	1.15±0.07
FS125	1.06±0.06
FS225	0.58±0.08
FS055	1.07±0.07
FS155	0.93±0.07
FS255	1.38±0.08
FS193	1.08±0.05
FS293	1.16±0.05

Table 7.2: The scaling factors calculated for 14 subruns in data, with each subrun run using a different SMELLIE fibre. All the subruns used the superK laser at 495nm. The errors are statistical errors only.

statistical errors. FS225 is clearly an outlier here, with a scaling factor of 0.58 ± 0.08. However, it seems likely that this is related once again to rope positions, as it is clear from Fig. 6.6c that a rope passes extremely close to the centre of the beamspot. The impact of ropes on results from this fibre is explored in Section 7.4.3.

Once a scaling factor has been calculated for each of the fibres at a single wavelength, the scaling factors are combined by calculating an error weighted mean. An error weighted mean is used so that outlying subruns with low statistics contribute less.

At 495 nm, using the superK, the scaling factor measured across all fibres is 1.10 ± 0.16. This error is the error weighted standard deviation, chosen to represent the full range of scaling factors calculated across the 14 fibres.

7.1.3 Wavelength Dependent Analysis

Once the scaling factor across all fibres has been measured at a single wavelength, attention turns to a wavelength dependent measurement.

Using the wavelength dependence of the measured scaling factor, the nature of the scattering can be investigated. The model of the scattering in RAT, as previously discussed, is purely Rayleigh scattering. Therefore, if a deviation from a flat scaling factor across wavelength is observed in data, the model does not completely describe

the scattering processes in the detector. An example could be the presence of Mie scattering in the detector.

To do a wavelength dependent measurement, the procedure for a finding scaling factor across all fibres is repeated at several wavelengths. The conservative region is re-found for every wavelength. Then, a flat line (of the form $y = $ p0) is fitted to the mean scaling factors versus wavelength.

The wavelengths chosen for this analysis are 375, 435, 495 and 555 nm. The subruns at 375 nm are performed using the PQ375 laser, as the wavelength range of the superK laser does not extend this short. However, the subruns for the other three wavelengths used the superK laser. This is primarily due to the superior stability of the superK laser, as discussed in Chapter 5. 375 nm was picked as one of the wavelengths as this is the shortest wavelength available to SMELLIE. Ideally, the longest wavelength possible would also be used. While the superK has a wavelength range that extends past 555 nm, the interest in photons above this wavelength range is limited, due to the low and highly variable PMT to PMT efficiency of the SNO+ PMTs in the wavelength region longer than this (see Chapter 2), the low emission probability from PPO and bisMSB (see [87]), and the small fraction of Cherenkov light above this wavelength. Accurately simulating subruns at wavelengths longer than 555 nm is also challenging due to the extraordinarily large number of photons which have to be produced and tracked to produce a single hit.

Until this point in the analysis, all the data used has been from a single SMELLIE run (114018). This is to ensure as consistent a detector state as possible across all the subruns. However, for operational reasons, SMELLIE runs were split into wavelength regions - for example, one run for wavelengths up to 495 nm. Therefore, it is not possible to have subruns at all the desired wavelengths in the same run. Therefore, when picking the subruns to use, runs were chosen within the same data taking period, meaning that they were all taken within eight hours of each other (during a detector shift). The runs used for this full scattering analysis are 114018, 114023 and 114034.

In Fig. 7.15, the scaling factor for every fibre (except FS093) at 375, 435, 495 and 555 nm are shown, with their statistical errors. In this plot, marker shapes are used to denote the fibre pointing angles (circles for 0° fibres, squares for 10° and triangles for 20°). Following this, in Fig. 7.16, the error weighted means and error weighted standard deviations, as error bars, are shown for the same wavelengths.

Unsurprisingly, some wavelengths provide less reliable scaling factor measurements than others. At 555 nm, the PMT efficiency is over two times lower than at its peak (at 439 nm). Compounding this, the absorption length of water is approximately
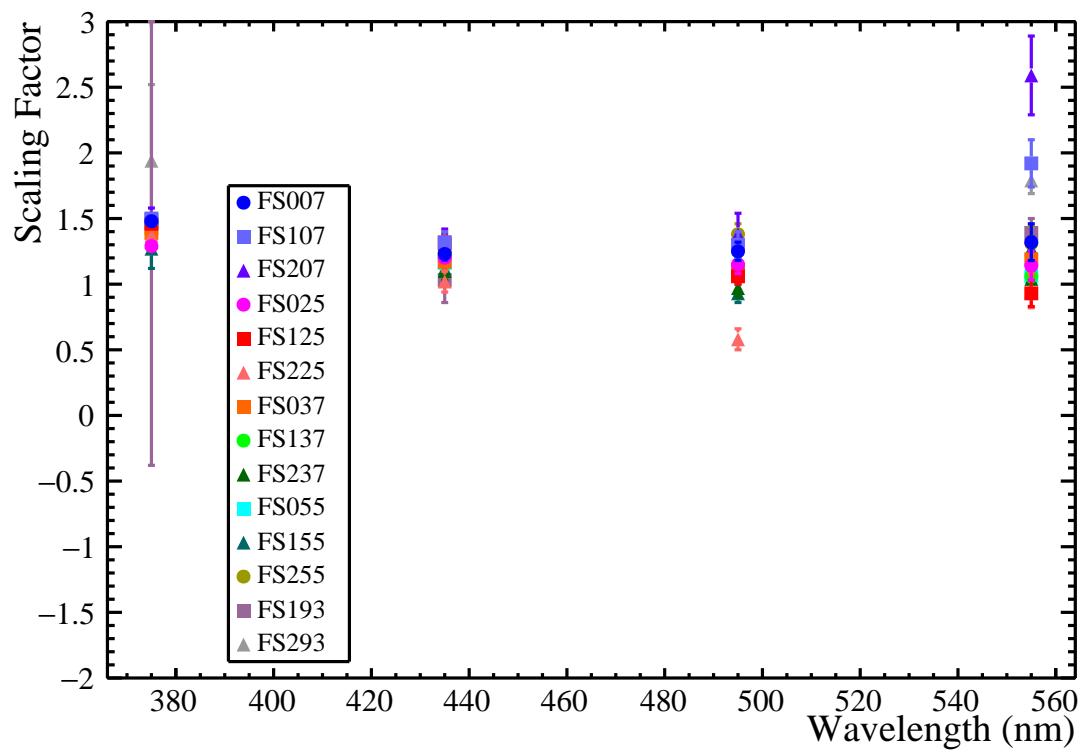
Figure 7.15: Scaling factors calculated for each subrun (for each fibre at 375, 435, 495 and 555 nm) with statistical errors. The different marker shapes denote the pointing angles of the fibres: circles for 0° fibres, squares for 10° and triangles for 20°.
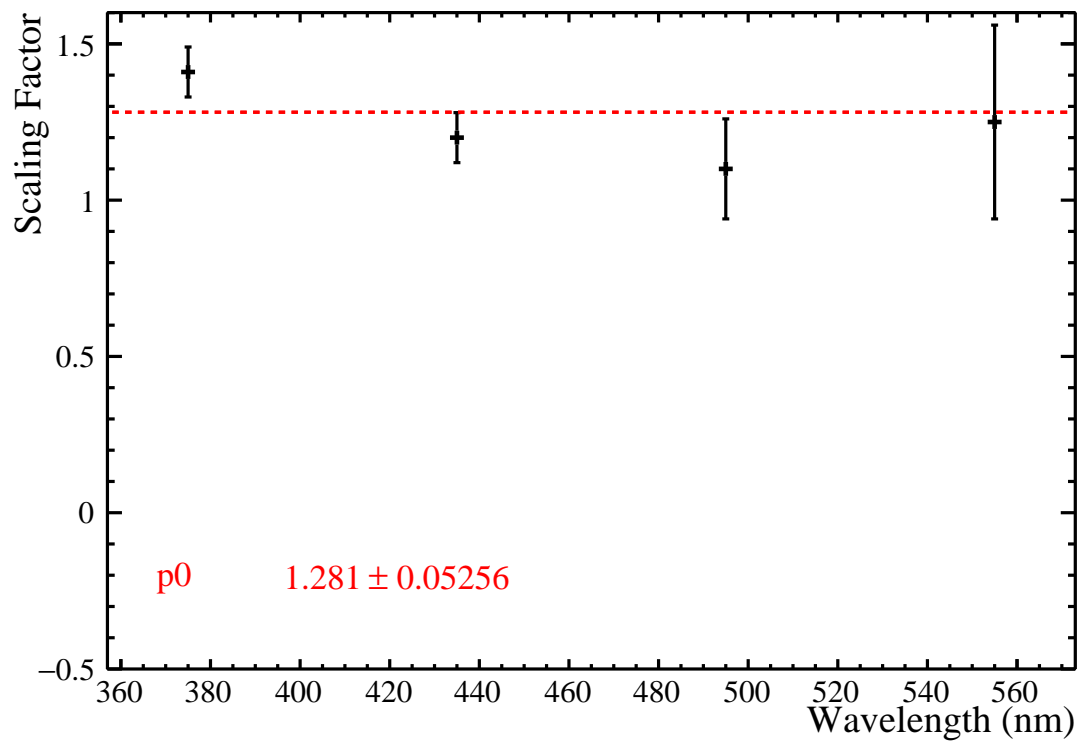
Figure 7.16: Error weighted means of scaling factors calculated for each wavelength with the error weighted standard deviations as error bars. The means are fitted with a function of the form $y = \text{p0}$. The result of the fit is shown in the bottom left of the plot.

20 times longer at 439 nm and the scattering length of water is approximately three times shorter. This means that at 555 nm, approximately 6.6 times as many photons are absorbed while 0.96 times as many photons are scattered. Therefore, there are fewer scattered photons as a proportion of the total number that hit PMTs, causing an increase in the statistical error on the measured scaling factor compared to 439 nm. This is evident in Fig. 7.15.

Despite this, the scaling factors measured as a function of wavelength are consistent with a constant. This is shown in Fig. 7.16, in which a fit of the form $y = $ p0 has been applied. In this case, p0 is fitted as 1.28 ± 0.05.

It is worth pointing out that one of these subruns has an extremely large error bar in comparison to the others. This is the subrun associated with FS193 using the PQ375 laser. The reason for this is extremely low statistics in this particular subrun. Ideally, a different subrun of FS193 with PQ375 would be used. However, the SMELLIE runs were taken in shifts that were months apart, meaning that using a single subrun from a different period of data taking would involve a high likelihood of a different detector state, making comparison difficult.

In conclusion, the scattering observed in run 114018 is consistent with Rayleigh scattering alone.

7.2 Scattering Length Analysis with Proposed Optics Constants

As discussed in Chapter 3, during the course of the above analysis and following analysis of laserball data, a new set of optical constants was proposed. These are referred to in this work as the proposed optics and are discussed in detail in Section 3.2.2.

Due to this change in the optics constants, the SMELLIE wavelength dependent scattering analysis was performed again, this time using the proposed optics. This was to both investigate whether the proposed optics made a difference to the scaling factor measured for the internal water and, simultaneously, to see if they were a viable set of optics constants for the detector from the point of view of SMELLIE. The acrylic scattering length change was of particular interest as this was the only change attributed to scattering. Other than the different optics constants, the same version of RAT and same data subruns were used as in the original analysis. This was, as before, done using 14 of the 15 fibres (FS093 is excluded) and with the same wavelengths: 375 nm using a PQ laser and 435, 495 and 555 nm using the superK laser.

Fibre	Photon Number	Percentage Difference (%)
FS007	1020 ± 9	-20
FS107	921 ± 8	-21
FS207	867 ± 8	-25
FS037	948 ± 8	-20
FS137	1109 ± 10	-21
FS237	883 ± 8	-20
FS025	796 ± 8	-19
FS125	947 ± 9	-14
FS225	954 ± 8	-21
FS055	833 ± 7	-19
FS155	990 ± 9	-21
FS255	729 ± 7	-19
FS193	3062 ± 20	-22
FS293	2868 ± 23	-24

Table 7.3: The photon numbers calculated for fourteen subruns in data, with each subrun using a different SMELLIE fibre. These have been calculated using the same method and data as in Table 7.1 except that the simulations have been run using the proposed optics constants. The errors are statistical. Also shown is the percentage difference between these numbers and the equivalent ones in Table 7.1.

Changing absorption lengths means that the number of photons to fire into the detector per shot for each subrun has to be recalculated. Table 7.3 shows the resultant photon numbers for the subruns at 495 nm, as well as the percentage difference from Table 7.1. This demonstrates the difference that a set of optics constants can make. On average, each subrun has been calculated to have approximately 20% fewer photons emitted per shot using the proposed optics than with the RAT optics. A decrease in the photon number is expected as, although the attenuation length in the acrylic is decreased (impacting both the scattering and absorption length), the absorption length in both internal and external water has increased. This means that, overall, fewer photons are absorbed and therefore fewer photons need to be fired into the detector to produce the observed number of hits. As a very simple estimate, the change in absorption length from the external water alone results in approximately 0.78 times as many photons absorbed in the simulations with the proposed optics (at 500 nm).

The effect that the optics changes have in the SMELLIE analysis is summarised in Fig. 7.17. This is the proposed optics equivalent of Fig. 7.2 and has been made using the same fibre and wavelength (FS007 at 495 nm using the superK). The different categories of optical effects are shown individually in Figs. 7.18 - 7.25.
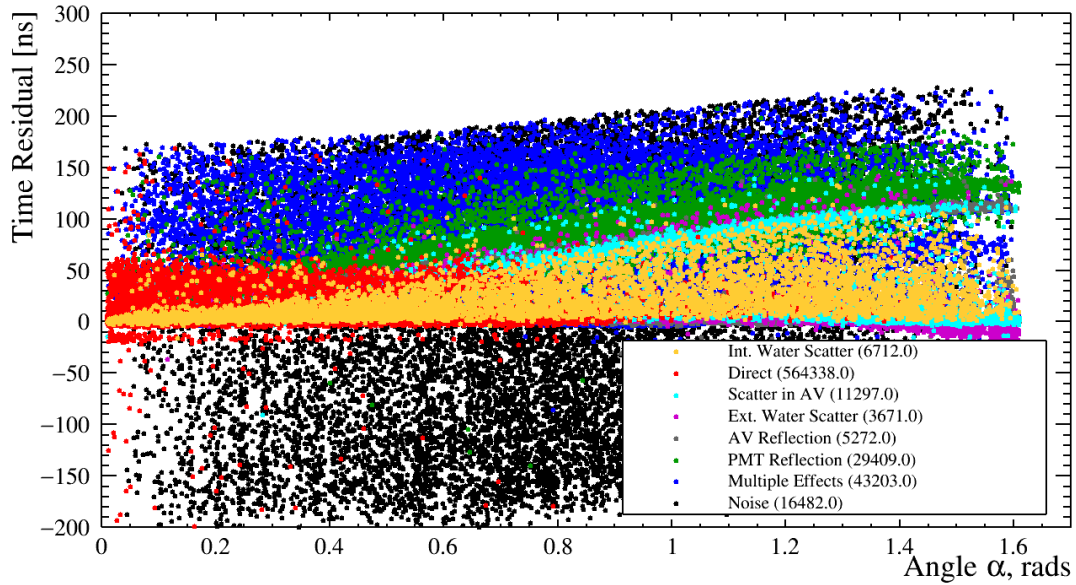
Figure 7.17: Time residual versus α angle for photons in a simulated SMELLIE subrun broken down into optical categories. This plot was created using the proposed optics constants, but uses the same fibre, wavelength and method as Fig. 7.2.
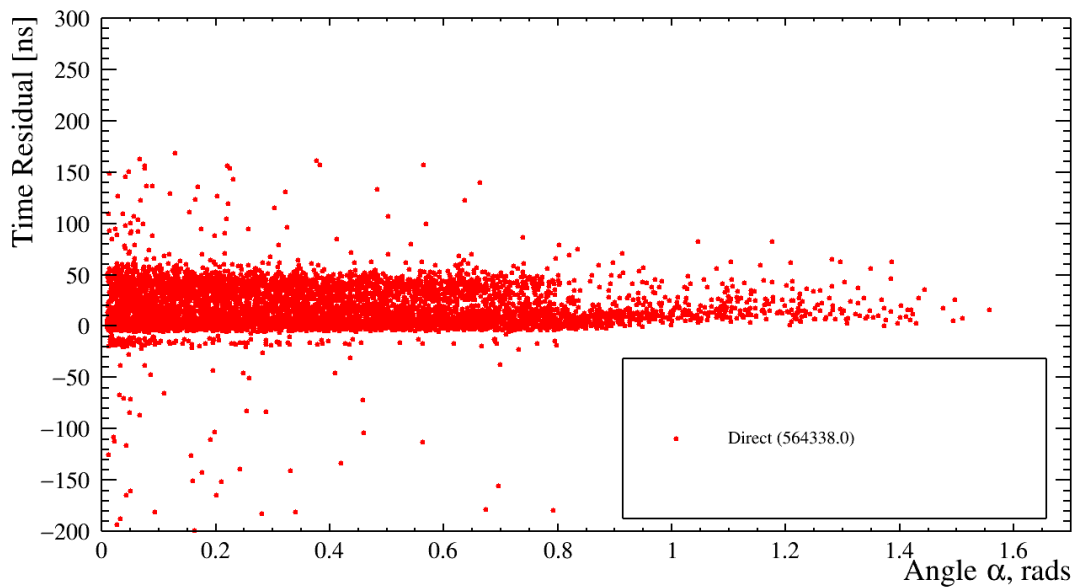
Figure 7.18: Time residual versus α angle for photons having reached PMTs directly in the simulated SMELLIE subrun shown in Fig. 7.17.
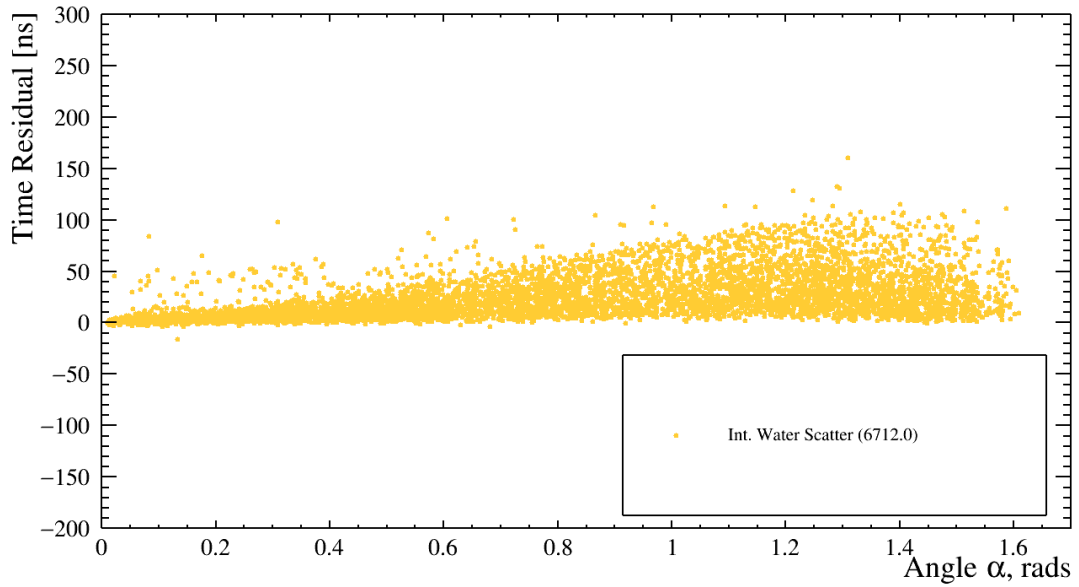
Figure 7.19: Time residual versus α angle for photons singly scattered in the internal water in the simulated SMELLIE subrun shown in Fig. 7.17.
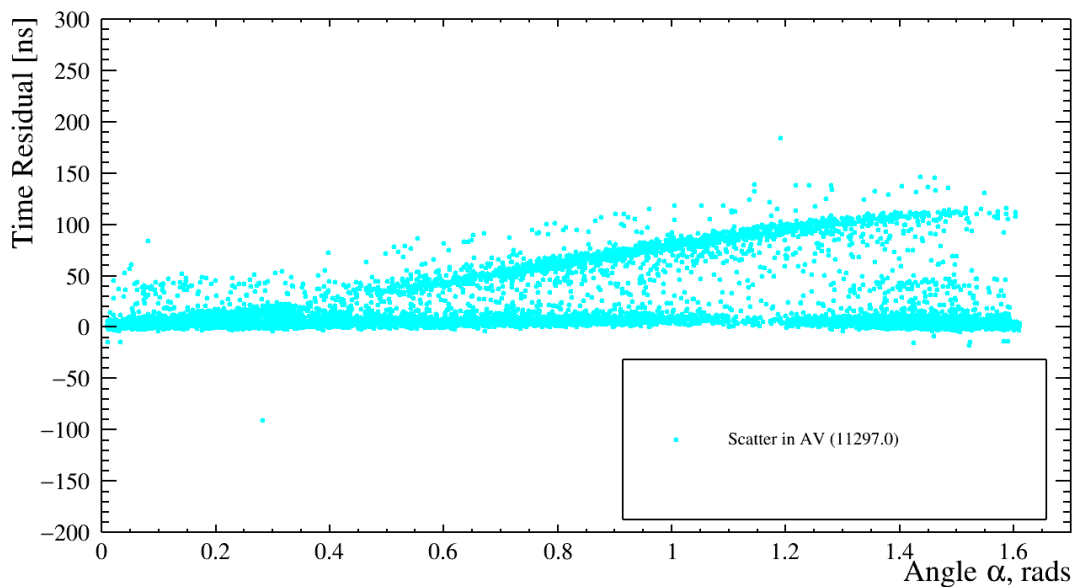
Figure 7.20: Time residual versus α angle for photons singly scattered in the acrylic in the simulated SMELLIE subrun shown in Fig. 7.17.
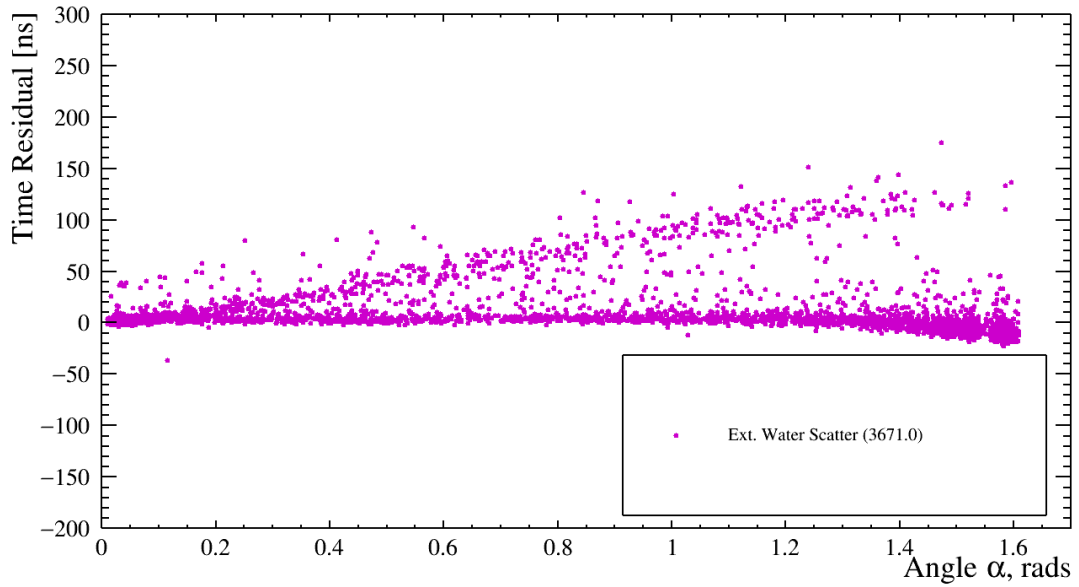
Figure 7.21: Time residual versus α angle for photons singly scattered in the external water in the simulated SMELLIE subrun shown in Fig. 7.17.
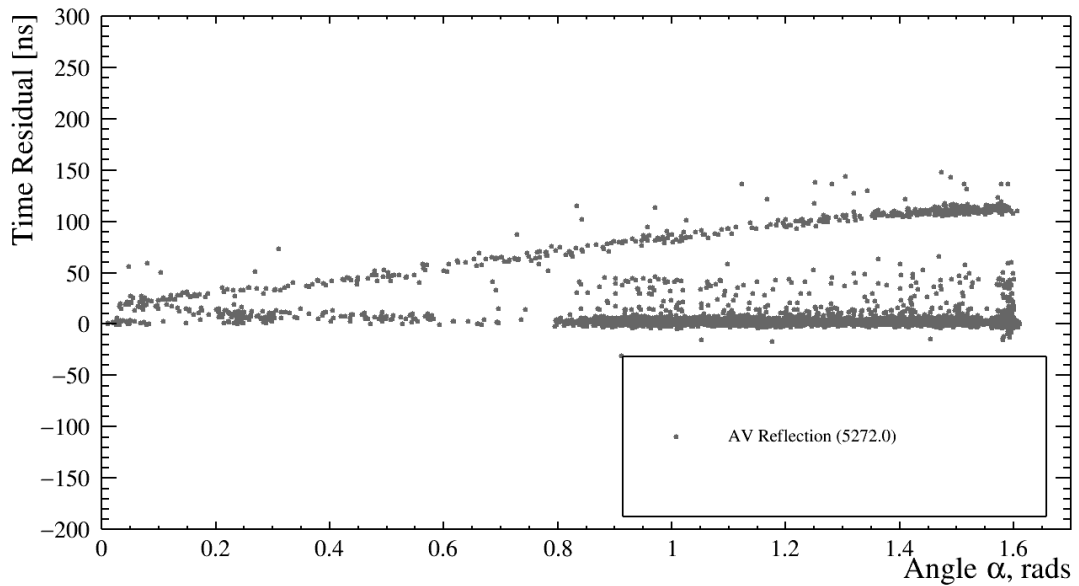
Figure 7.22: Time residual versus α angle for photons reflected by an AV/water interface in the simulated SMELLIE subrun shown in Fig. 7.17.
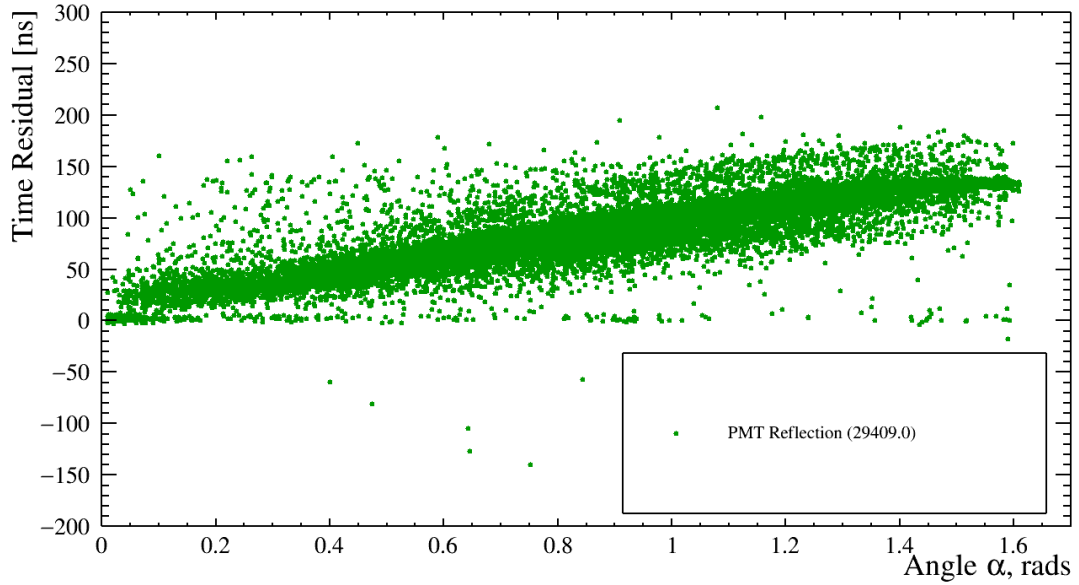
Figure 7.23: Time residual versus α angle for photons reflected from a PMT in the simulated SMELLIE subrun shown in Fig. 7.17.
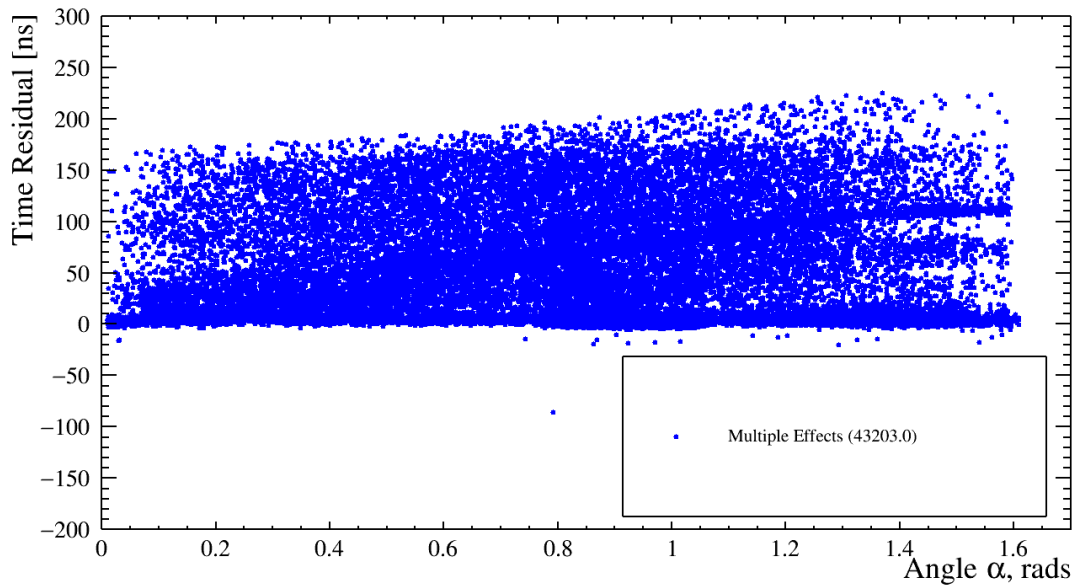
Figure 7.24: Time residual versus α angle for photons having undergone more than one optical effect in the simulated SMELLIE subrun shown in Fig. 7.17.
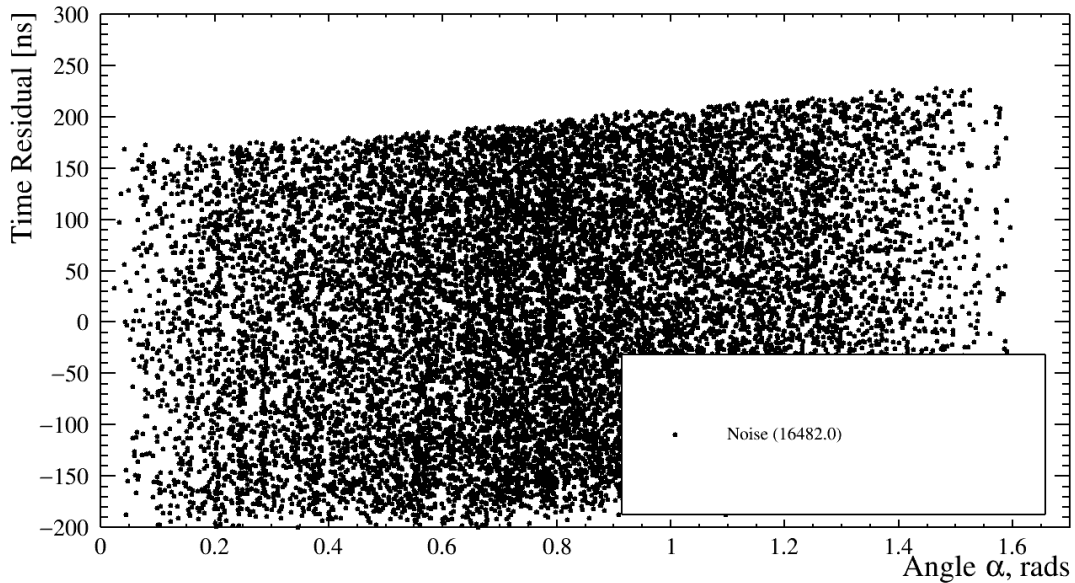
Figure 7.25: Time residual versus α angle for hits caused by detector noise in the simulated SMELLIE subrun shown in Fig. 7.17.

As can be seen in Figs. 7.5 and 7.20, with the proposed optics there is, as expected, approximately 100 times more scattering occurring in the acrylic. There are 130 photons scattered once in the AV in the original optics compared to 11297 in the proposed optics. In the original optics, the probability of a photon passing through 10 cm of acrylic without being scattered at 500 nm (ignoring absorption) was 0.999 (to 3 s.f.). However, in the new optics, this probability has been reduced to 0.946.

The optimisation of the selection region has been redone for these proposed optics, rather than reusing the region from the original optics. As before, the selection region was found for each fibre and then the most conservative combination of gradients and intercepts was chosen so that the same region was used for all fibres at each wavelength.

The percentage contributions of each optical category to the selection region for both versions of the optics are shown in Table 7.4. While the contribution of AV scattered photons in the selection region has increased by approximately 5%, the signal in the selection region using the proposed optics has decreased by approximately 12%. However, the largest effect is for photons undergoing multiple optical interactions. The contribution of this category increases by approximately 15% using the proposed optics. This is understandable - the increase in scattering in the acrylic increases the probability that photons will be detected after undergoing both an AV scatter and another optical effect.

Category	Contribution - Original (%)	Contribution - Proposed (%)
Direct	11.5 ± 0.5	7.9 ± 0.4
Int. Water Scatter	44.2 ± 1.1	32.4 ± 0.8
AV Scatter	0.2 ± 0.1	5.3 ± 0.3
Ext. Water Scatter	2.8 ± 0.2	2.1 ± 0.2
AV Reflection	2.9 ± 0.2	2.6 ± 0.2
PMT Reflection	0.6 ± 0.1	0.7 ± 0.1
Multiple Effects	21.4 ± 0.7	36.3 ± 0.9
Detector Noise	16.3 ± 0.6	12.8 ± 0.5

Table 7.4: The percentage contributions to the selection region (to 1 d.p.) from each category of optical effect in both the original and proposed optics.

The expectation was that this change would result in a decrease in the scaling factor measured for a given data subrun, as hits in the selection region are increasingly modelled as originating from other optical categories. Therefore a smaller amount of scattering in the internal water is required for the observed number of hits in the selection region. A flat scaling factor as a function of wavelength is still expected. However, this was not observed.

The results for each subrun at each of the wavelengths are shown in Fig. 7.26. The error bars, as before, represent the statistical error. These results, compared to Fig. 7.15, are not flat with respect to wavelength. This is further evident when looking at the mean of the scaling factors calculated for each wavelength, shown in Fig. 7.27. For comparison purposes, this plot has been fitted with a function of the form $y = \mathrm{p0}$ and the result of the fit is shown in bottom left of the plot. Clearly, this fit function is not suitable.

Furthermore, some fibres produce unphysical scaling factors - a negative scaling factor means that it is not possible to make a measurement of the scattering length. It is evident, also, that as the wavelength is increased, more subruns result in negative scaling factors, or scaling factors consistent with zero. Even at the most strongly scattering short wavelength of 375 nm, one subrun produces a scaling factor which is consistent with a negative result (which is known to be due to low statistics). At 555 nm, eight of the subruns produce results consistent with negative scaling factors. This indicates that the increase in scattering in the acrylic is making it infeasible to measure the scattering length of the internal water.

To understand why this is the case, the model of scattering in the acrylic must be considered. As shown in Chapter 3, the acrylic is modelled as a Rayleigh scatterer in RAT, which means the scattering length follows a λ^4 distribution. The motivation
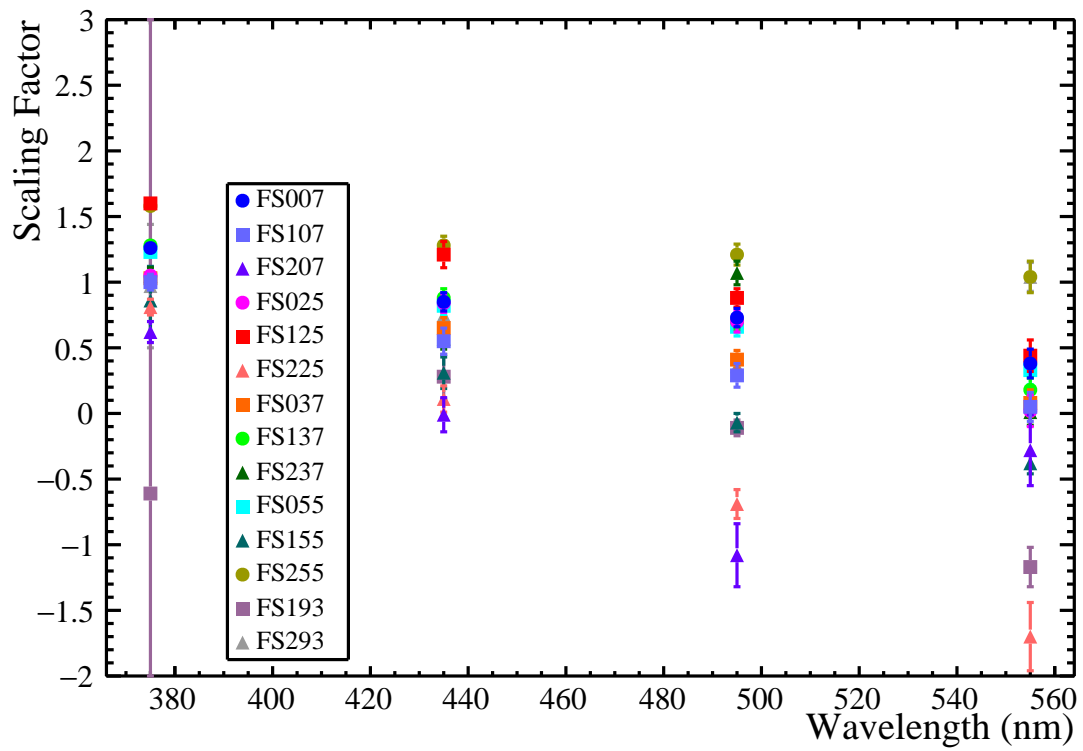
Figure 7.26: Scaling factors calculated for each subrun (for each fibre at 375, 435, 495 and 555 nm) with statistical error bars using the proposed optics. The different marker shapes denote the pointing angles of the fibres: circles for 0° fibres, squares for 10° and triangles for 20°.
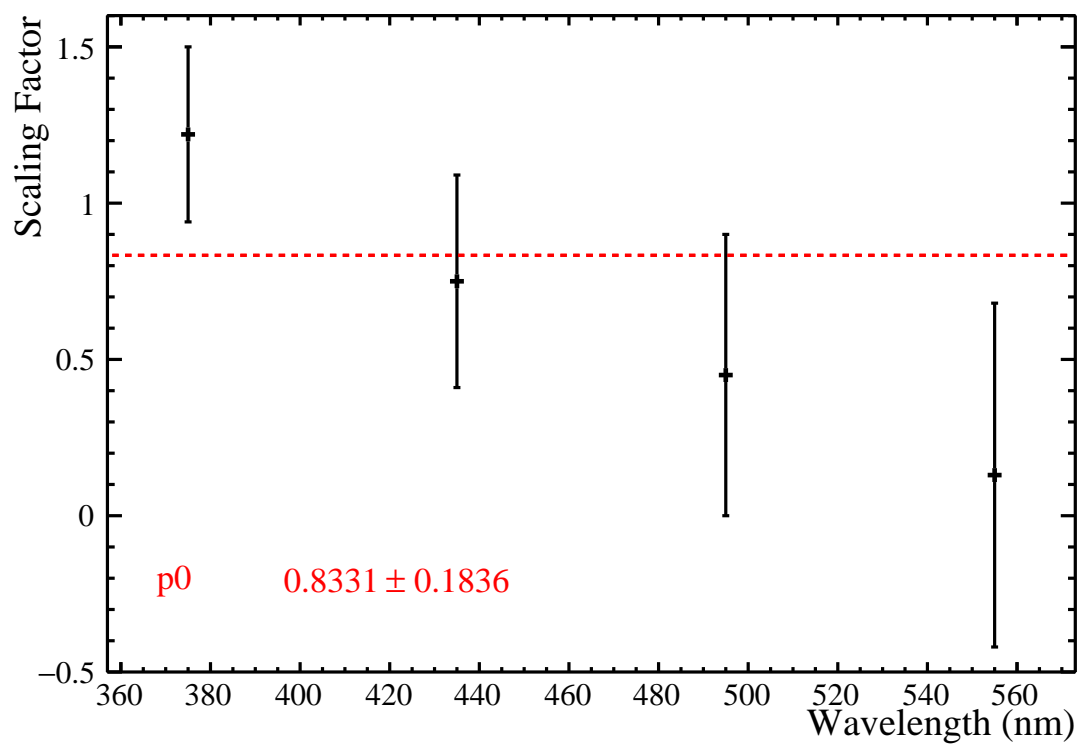
Figure 7.27: Error weighted means of scaling factors calculated for each wavelength with the error weighted standard deviations as error bars using the proposed oprics. The means are fitted with a function of the form $y = \mathrm{p0}$. The result of the fit is should in the bottom left of the plot.

for decreasing the scattering length of the acrylic by a factor of 100, as discussed in Section 3.2.2, was not that the bulk acrylic had become 100 times more likely to scatter, but that since SNO there had been an increase in surface effects, imperfections and crazing. Therefore, the additional scattering is not expected to be Rayleigh scattering.

As shown in Table 7.4, before the acrylic scattering length was decreased, scattering in the acrylic was a negligible effect and therefore the spectral distribution of that scattering was unimportant. However, with the proposed optics, acrylic scattering is no longer negligible and therefore modelling it as Rayleigh scattering is, as shown in this work, no longer consistent with the SMELLIE measurements.

Additional support for this conclusion arises from the correlation between the scaling factor measurement of each individual fibre and the pointing angle of that fibre. In Fig. 7.27, the fibre pointing angles are denoted by using different symbols as a markers (circles for 0° fibres, squares for 10° and triangles for 20°). 20° fibres, at 555 nm, are generally the ones producing the most unphysical scaling factors. Four out of the five 20° fibres result in a scaling factor that is either negative or compatible with zero. Three out of five 10° fibres do, but only one out of four 0° fibres. This supports the theory that the change in acrylic properties causes the unphysical measurements: the 20 degree fibres pass through the greatest length of acrylic. This is shown in Fig. 7.28, where the most likely path length in acrylic for FS007 is approximately 110 mm but approximately 121 mm for FS207.

In conclusion, the proposed acrylic scattering spectrum is not only incompatible with measurements of the scaling factor of internal water at longer wavelengths, it is stealing the wavelength dependence from the internal water. If Rayleigh scattering occurs this often in the acrylic, the data cannot be consistent with a λ^4 scattering length distribution in the internal water. It would be possible that this is true if there is another form of scattering that is dominant in the internal water. However, this is not expected in the water phase of SNO+, which utilises some of the world's purest water inside the AV, especially as the internal water attenuation length has been measured to be longer in SNO+ than the external water in SNO. Also, the result of the scattering analysis using the original optics results in internal water behaviour that is consistent with Rayleigh scattering. This adds additional weight to the expectation that there is not another scattering process dominant. Therefore, this is a strong indication that the proposed optics are incorrect.

While this is a demonstration that the scattering length in the acrylic is not 100 times shorter than previously thought, it does not contradict the laserball measure-
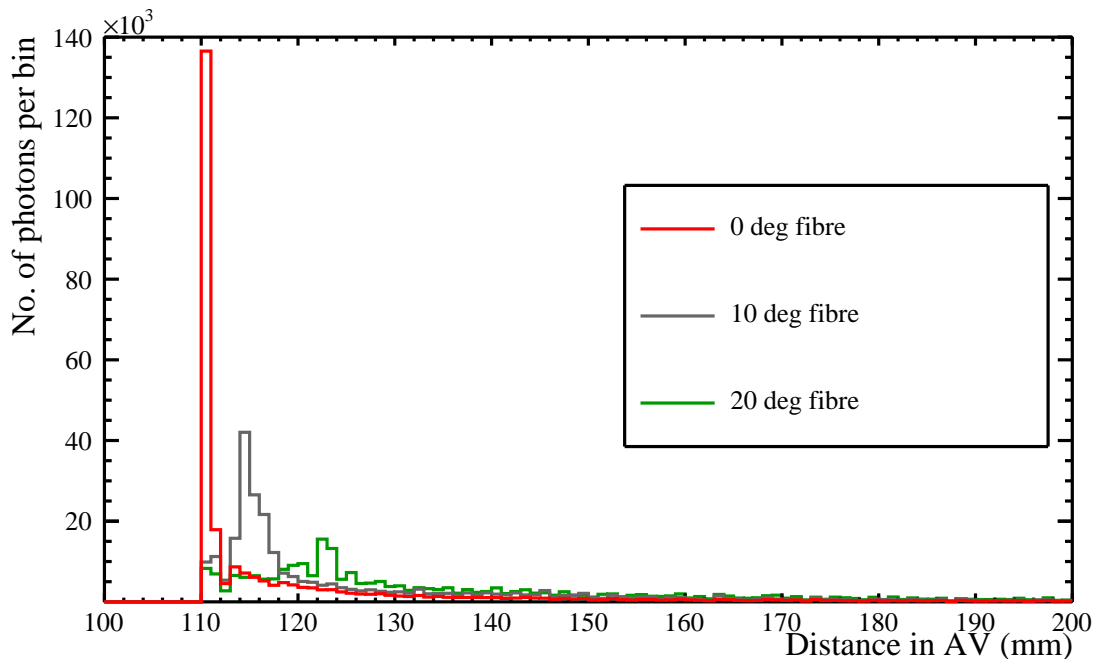
Figure 7.28: The distance travelled by each photon through the acrylic in simulations of fibres on the 07 node at the three different pointing angles using the superK at 495 nm.

ments of the attenuation length. This analysis does not make a statement on whether the acrylic attenuation has changed since SNO, only that the scattering component has not changed by a factor of 100.

7.3   A Scaling Factor for SNO+ Water Phase

Due to the previous result, another set of optical constants were calculated, referred to as the 'updated optics'. These are discussed in detail in Section 3.2.3. The changes involved with moving from the original RAT optics to the updated optics need to be implemented in the scattering length analysis. However, rather than rerunning the analysis with all the changes at once, the changes are introduced to the analysis one at a time, meaning that any difference that arises can be accurately and immediately identified and associated with the optical change which caused it. In addition, due to time constraints, it was not feasible to repeat the analysis with all fourteen fibres at all four wavelengths when each change is implemented. Therefore, the analysis is repeated for each change for two fibres (FS007 and FS237) at three wavelengths. FS007 and FS237 are chosen because they are a $0°$ and $20°$ fibre, representing the two extremes of pointing angle. These two fibres have been well behaved in the

previous analysis, not resulting in outlying results. They also do not have ropes passing through the centre of their profiles. The three wavelengths chosen are 375, 435 and 555 nm. Again, these include the two extremes of wavelength and also the shortest superK wavelength used in the previous analysis.

Figs. 7.29 - 7.32 show the changes introduced one at a time, as discussed above. Firstly, Fig. 7.29 shows the results of the original analysis, except that the external water scattering scaling factor has been varied identically to that of the internal water. Next, this analysis is taken and repeated with the additional change that the attenuation lengths for the proposed optics have been used for both the internal and external water. These results are shown in Fig. 7.30. Fig. 7.31 shows the results when, as well as the changes already discussed, the absorption lengths of the internal and external water have been coupled to their respective scattering lengths according to Eq. 3.2 in order to keep the attenuation length constant. Finally, Fig. 7.32 shows the results with all of the necessary changes (i.e. all of those discussed so far and the updated acrylic optics).

As can be seen from these plots, the scaling factor results are largely completely consistent with the original scaling factor analysis. At 375 nm, both fully updated results are within three standard deviations of the mean of the original analysis (and FS007 is within one standard deviation). This is also true at 435 nm. At 555 nm, one of the results is within two standard deviations of the mean of the original analysis and the other within three.

However, as can be seen from Fig. 7.32, at 555 nm, the scaling factors from the updated analysis appear noticeably lower than the original results and it is clear from the progression of introducing the changes that they get further from the original mean with each change that is introduced. On the other hand, as previously discussed, 555 nm provides the least reliable of the scaling factor measurements, evidenced by the error weighted standard deviation being almost four times larger at 555 nm than at 375 nm. In addition, 555 nm is, typically, a longer wavelength than those of interest in SNO+ physics analyses. This decrease should be investigated further but it was decided that in water phase it is not a concern. Therefore, the originally measured scaling factor for internal water is still valid, regardless of the optical changes introduced. This meant that the entire analysis did not need to be repeated.

In conclusion, a divisive scaling factor of 1.28 ± 0.05 (stat.) is applied to the scattering length of the updated internal and external waters in RAT and the absorption lengths are adjusted according to Eq. 3.2 to account for this. This is the final version of the optics constants for water phase.
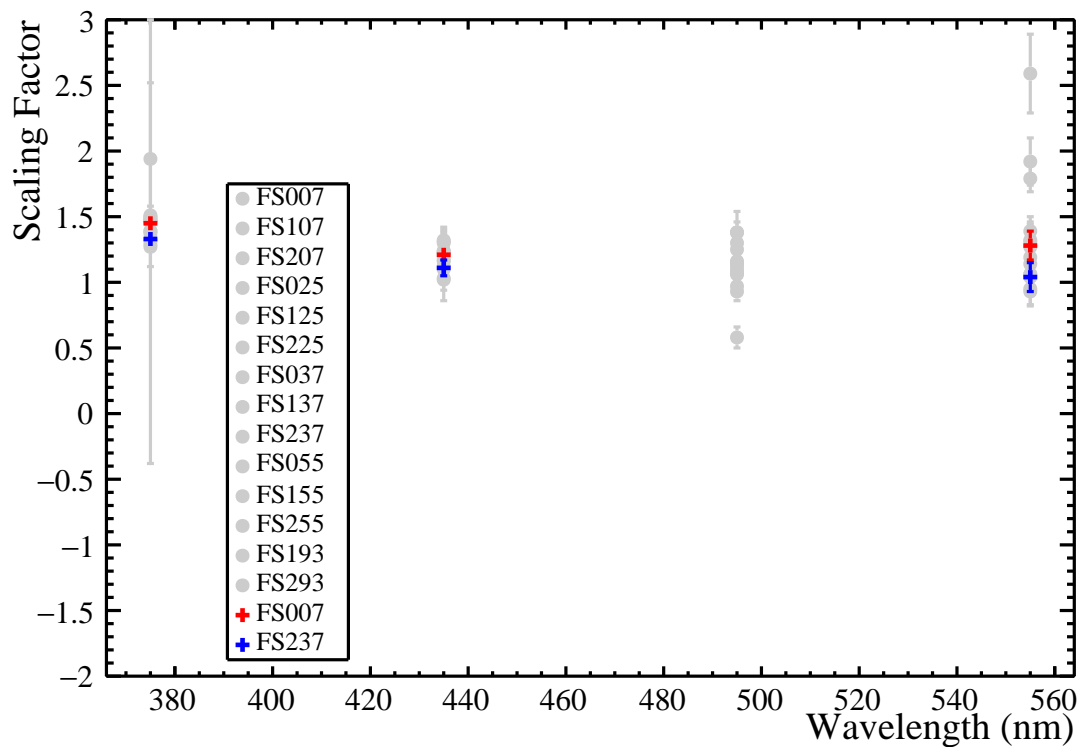
Figure 7.29: Testing the effect of varying the scattering length of the external water identically to the internal water in the scattering analysis, using the otherwise original RAT optics. The grey points are the original results of the scattering analysis for each fibre, using the RAT optics, and the two colour crosses (red for FS007 and blue for FS237) show the results of the two fibres at the three wavelengths used to test implementing the coupled scattering length of the internal and external water.
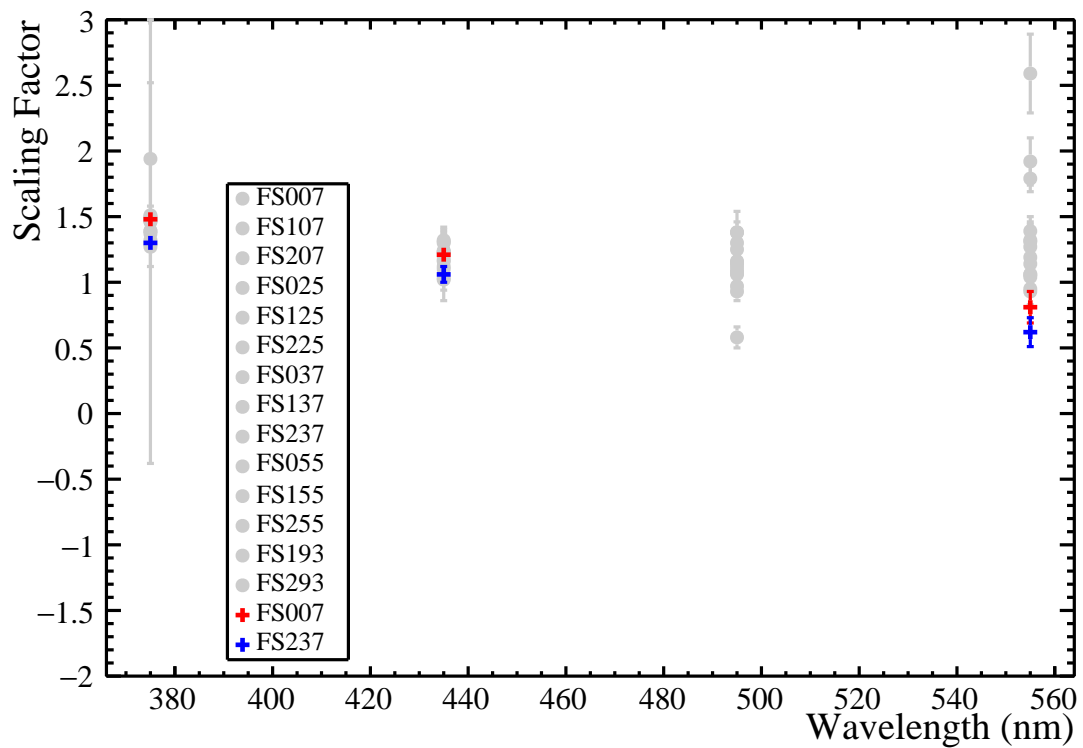
Figure 7.30: Testing the effect of using the updated water optics but the original RAT acrylic optics in the scattering analysis, as well as varying the scattering length of the external water identically to the internal water. The grey points are the original results of the scattering analysis, using the RAT optics, for each fibre and the two colour crosses (red for FS007 and blue for FS237) show the results of the two fibres at the three wavelengths used to test implementing the change to the optics.
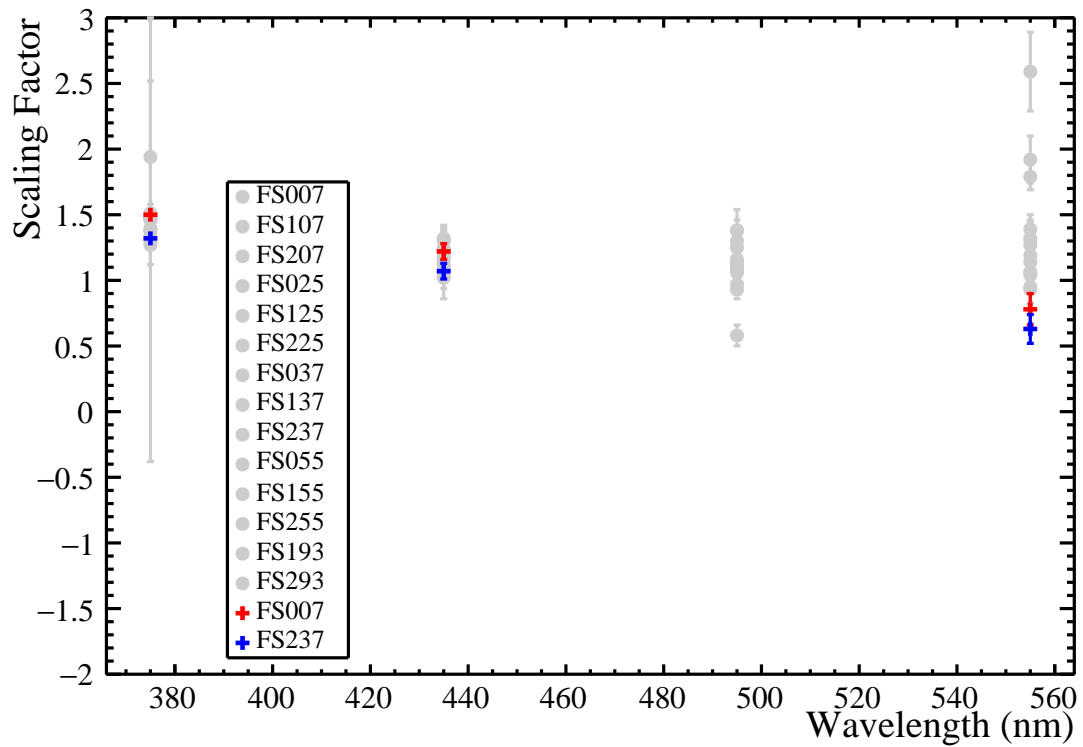
Figure 7.31: Testing the effect of keeping the attenuation lengths of internal and external water constant by varying the absorption lengths alongside the scattering lengths in the scattering analysis, as well as using the updated water optics but the original RAT acrylic optics and varying the scattering length of the external water identically to the internal water. The grey points are the original results of the scattering analysis, using the RAT optics, for each fibre and the two colour crosses (red for FS007 and blue for FS237) show the results of the two fibres at the three wavelengths used to test implementing the change to the optics.
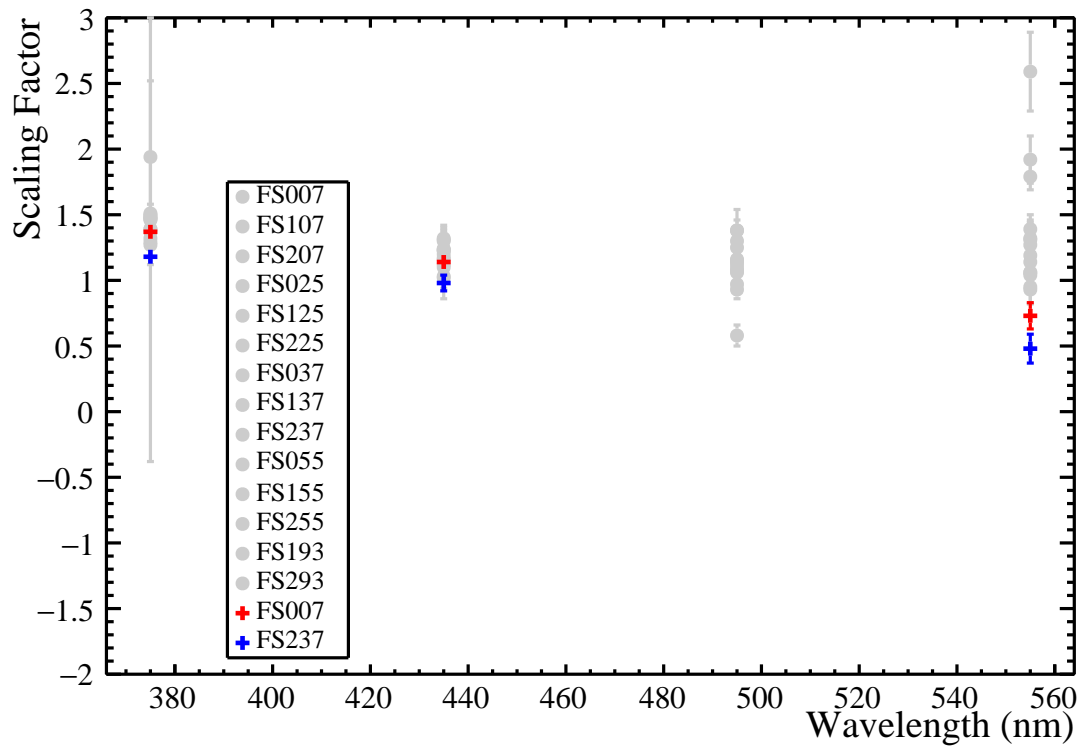
Figure 7.32: Testing the effect of using the updated acrylic optics in the scattering analysis as well as varying the absorption length alongside the scattering length in order to keep the attenuation length constant, using the updated water optics and also varying the scattering length of the external water identically to the internal water. The grey points are the original results of the scattering analysis, using the RAT optics, for each fibre and the two colour crosses (red for FS007 and blue for FS237) show the results of the two fibres at the three wavelengths used to test implementing the change to the optics.

7.4 Systematic Errors

After determining the UPW scattering scaling factor, the next step is to measure the systematic error on this measurement. All the errors that have been given so far have been statistical errors.

As the scattering analysis is a Monte Carlo driven analysis, it is crucial for the simulation to be an accurate representation of the data for the result to be meaningful. Therefore, the majority of the systematic errors in this analysis arise from considering the possible cases in which the simulation deviates from reality.

This analysis is designed to be systematics dominated. The number of shots taken in each SMELLIE subrun was chosen to provide adequate statistics such that statistical errors should be negligible compared to the systematics involved. If this was not the case, subruns of SMELLIE data taken with the same fibre at the same wavelength and at similar intensities could be combined.

Following is a discussion of the various systematic errors which are expected to contribute the most to the overall error in this analysis.

7.4.1 Beam Normalisation Errors

The number of photons fired into the detector in the scaling factor simulations, as discussed in Section 7.1.1, is found through matching the number of p.e. per shot in the entire detector in data and simulation. However, to do this, assumptions have to be made concerning the scattering and absorption in the data. The number of p.e. seen in any given event has to weakly depend on the scattering length. This is true as scattering does not directly change the number of photons in the detector; it just shifts and smears the time distribution of p.e.. A shorter scattering length can cause a slight decrease in p.e. though, as an increased number of photons may not hit a PMT before the end of the event window due to multiple scatters. This is investigated in [53]. However, the number of p.e. has a much stronger dependence on the absorption length, as this directly removes photons from the detector.

Therefore, a mismodelling of the water absorption length (water is considered as the path lengths are orders of magnitude longer in water than in any other detector media) would cause a different result in the stage of analysis which calculates the photon number to use in the scaling factor simulations.

The systematic associated with this photon number normalisation is investigated in this section. To do this, the original optics in RAT and the superK at 495 nm have been used. The same fibres that were used to test optics changes have been used

again (FS007 and FS237). The errors on the attenuation length as measured by the laserball have been attributed purely to uncertainty on the absorption length. The photon number simulations have then been rerun within this uncertainty to create long and short extremes of the absorption length at 495 nm. In both cases, a photon number for the scaling factor simulations has been calculated.

The two extremes of photon number calculated are 875 and 1796 compared to 1271 found originally for FS007 and 759 and 1568 compared to 1107 for FS237. This is a variation of -31% to +41% or 42% for both fibres. This provides a good estimate of the largest possible error expected on the photon number.

For each of the fibres, the scaling factor simulations have then been run with these two extreme photon numbers and scaling factors calculated. These are compared to the scaling factor resulting from the original analysis. For FS007, the scaling factor measured by the original analysis is 1.25 ± 0.07. The scaling factors that result from this variation in photon number are 1.23 ± 0.09 and 1.17 ± 0.06 respectively. The results for the FS237 analysis are similar: a scaling factor of 0.97 ± 0.07 for the original analysis and 0.92 ± 0.09 and 0.93 ± 0.06 respectively for the analyses with varied photon numbers.

For both of the fibres, for both photon number extremes, the resultant scaling factors are statistically consistent with the originally measured scaling factor. Therefore, any systematic error resulting from photon number errors due to attenuation mismodelling is smaller than the statistical error and not possible to distinguish without using more data. Therefore, this is not a systematic error to be concerned about and does not need further investigation.

7.4.2 Beam Profiling Mismatch

As shown in Section 6.2.2, there remains a mismatch between simulation produced using the 2D beam profiles and data. The level of mismatch is shown as a function of α for FS055, FS055 and FS237 in Figs. 6.10 - 6.12. The χ^2_{DOF} between the data and simulation histograms are 4.51, 6.81 and 6.94, respectively.

The systematic introduced in the scaling factor analysis by this mismatch is evaluated in this section. It is not possible to create a beam profile without this mismatch between data and simulation. Therefore, running the scattering analysis without this mismatch is not possible, even using simulation. Therefore, an alternative method of evaluating the systematic error introduced by this mismatch must be found.

This is done firstly by creating a beam profile from detector data (for FS055 at 495 nm, using the superK), as normal, then simulating a fake data set using that
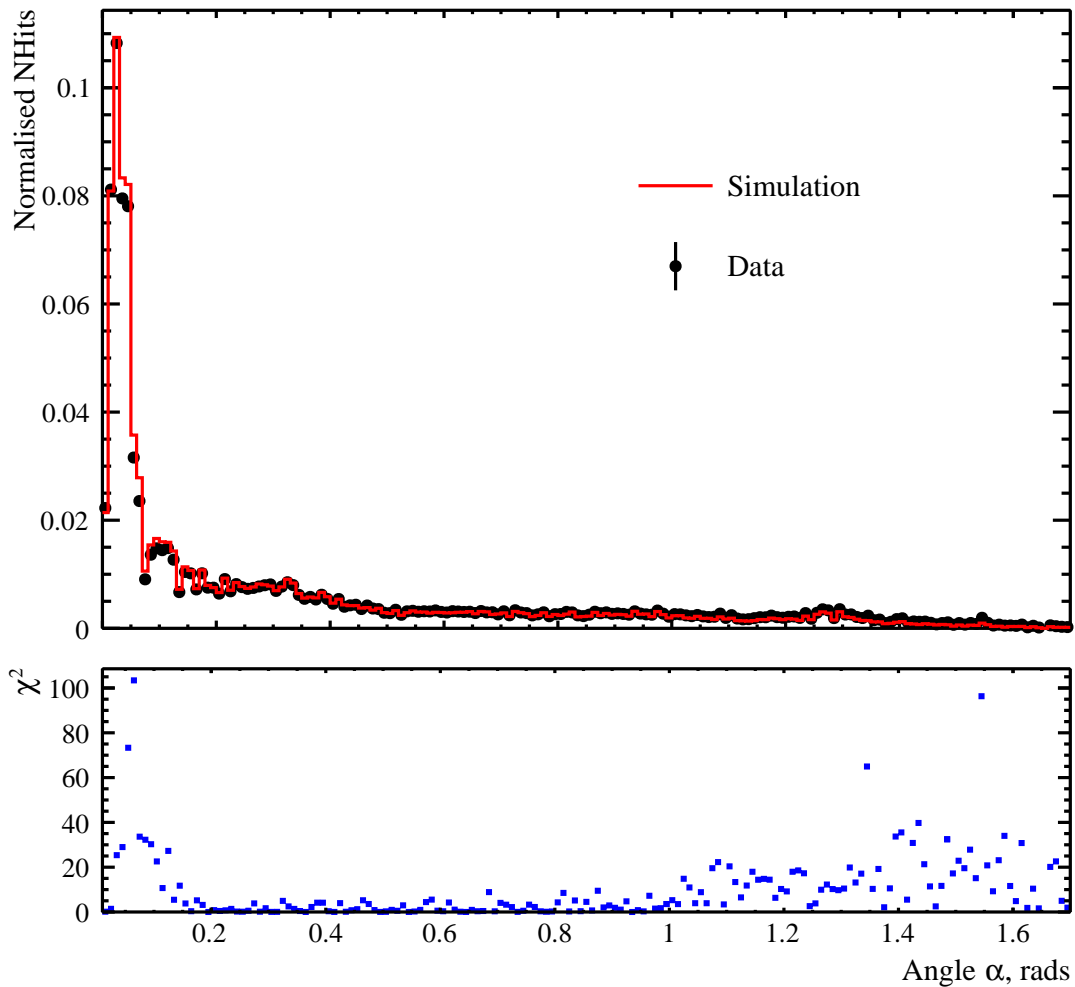
Figure 7.33: Simulation created using the fake data 2D beam profile of FS055 at 495 nm is shown in red, compared to the aforementioned fake data in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 10.05.

same beam profile. This fake data set is now treated as if it were real detector data, with the exception that the scaling factor used to create it is known. A beam profile is then created from this fake data set - a fake data beam profile. A further simulation is created using the fake data beam profile in order to compare the mismatch between this new simulation (for ease, called a fake simulation) and the fake data to the mismatch between the real data and a simulation based on a beam profile derived from the real data (such as shown in Fig. 6.10). The fake simulation and the fake data are shown in Fig. 7.33. In this case, the χ^2_{DOF} is 10.05.

Using Fig. 7.33, it is important to note that the mismatch between simulation and data (albeit both fake) shows the same characteristics as it did in Fig. 6.10; the

simulation overestimates the NHits on the falling edge of the beam (at approximately 0.05 rads).

As the scaling factor used to create the original fake data is known, the scattering analysis is performed on this fake data in order to evaluate the systematic error introduced by such a mismatch. This analysis resulted in a scaling factor that is completely statistically consistent with the scaling factor used to create the fake dataset. Therefore, the level of mismatch seen in the beam profiles in Section 6.2.2 causes no discernible systematic error in this analysis.

It is crucial to remember that in the above discussion the scattering analysis has been performed at the same wavelength the beam profile creation used. Ideally, if one is successful at isolating direct beam hits, the beam profiles should be wavelength independent. However, it is possible that light propagation through the fibre and collimator over the 200 nm wavelength range of the analysis causes a difference in the angular distribution of emitted light. Also, and less hypothetically, it is not possible to completely isolate direct beam hits using only time residual and angular cuts. There is no way to remove small angle scatters (which is the dominate effect of Rayleigh scattering) from the direct beam. Therefore, the aim should be to create beam profiles at the longest wavelength used in the scattering analysis as this should contain the least contamination from scattered light (due to the λ^4 dependence of the Rayleigh scattering length). However, the beam profiles used in this analysis were produced at 495 nm, which is not the longest wavelength examined. (This is partially due to the increase in photon number needed to create sufficient isotropic simulation at 555 nm and partially due to the vast majority of SMELLIE commissioning being performed at 495 nm.)

The effect of creating beam profiles at 495 nm and then using them to create simulation at a different wavelength is shown in Figs. 7.34 and 7.35, where simulation using the profiles has been created at 375 nm and compared to the corresponding 375 nm data subruns used for the scattering analysis. It is evident that there is a slight excess of NHits at angles just past the peak of the beam (approximately at 0.2 rads). The χ^2_{DOF} are significantly increased from Figs. 6.11 and 6.12 and are now 57.25 for FS007 and 15.66 for FS237. Further investigation into this worse match at wavelengths other than 495 nm should be done, but was unfortunately not possible within the time frame of this work.

The systematic error caused by using beam profiles created at 495 nm for the scattering analysis at three other wavelengths needs to be evaluated. In order to do this, new 2D beam profiles are created from data at 375 nm (using the PQ375) for

Figure 7.34: Simulation created at 375 nm using the 2D beam profile of FS007 made at 495 nm is shown in red, compared to the matching 375 nm data subrun in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 57.25.

Figure 7.35: Simulation created at 375 nm using the 2D beam profile of FS237 made at 495 nm is shown in red, compared to the matching 375 nm data subrun in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 15.66.

FS007 and FS237. 375 nm was chosen as this is the wavelength used in the scattering analysis that is the furthest from 495 nm, the wavelength at which the beam profiles were originally derived. Simulation produced using the profiles created at 375 nm are shown in Figs. 7.36 and 7.37, compared to the corresponding data subruns. The χ^2_{DOF} for these plots are 7.08 and 11.35, respectively, which are significant decreases from the χ^2_{DOF} calculated when using profiles created at 495 nm. It is clear that NHits in the region around 0.2 rads are a much better fit to data than previously.

Using the beam profiles created at 375 nm, the scattering analysis is repeated for FS007 and FS237. This results in scaling factors of 1.64 ± 0.03 and 1.44 ± 0.05 respectively. These are compared to the results of the original analysis: 1.48 ± 0.03 and 1.32 ± 0.04, respectively. For both fibres, the two measured scaling factors are not statistically consistent: using a beam profile created at 375 nm rather than 495 nm has caused a systematic effect on the measured scaling factors. This effect takes the form of an increase in the scaling factor.

When directly comparing the two profiles, it is clear that the profile made at 375 nm has, relatively, fewer hits in the central beam, an increase of hits in the region around 0.2 rads (where there was previously a deficit compared to the data) and then fewer hits again above angles of approximately 0.6 rads. It is this high angle region which is of most interest in the scattering analysis: an decrease of hits here results in a lower fractional acceptance for the simulations and, therefore, a higher scaling factor calculated for data.

This results in a systematic of 11% for FS007 and 9% for FS237. It is expected that FS007 would have a larger systematic error, as the difference in χ^2_{DOF} between Figs. 7.34 and 7.36 (for FS007) is over 11 times larger than for FS237. Therefore, in order to be conservative, the systematic error that arises from mismatch between simulation and data due to the beam profiles is taken to be 11%.

7.4.3   Mismodelling of Ropes in the Beam Profile

The systematic error from a mismatch between data and simulation due to the wavelength at which the beam profiles have been extracted has already been discussed. However, there is another, more specific, cause of inaccuracy in simulation which arises from the beam profile. As was discussed in Chapter 6, there is a difference in the location of the rope shadows in data and simulation. This was noticed when creating the 2D beam profiles for each of the fibres. From SMELLIE data alone, correcting the rope positions would be challenging, as the SMELLIE beams do not
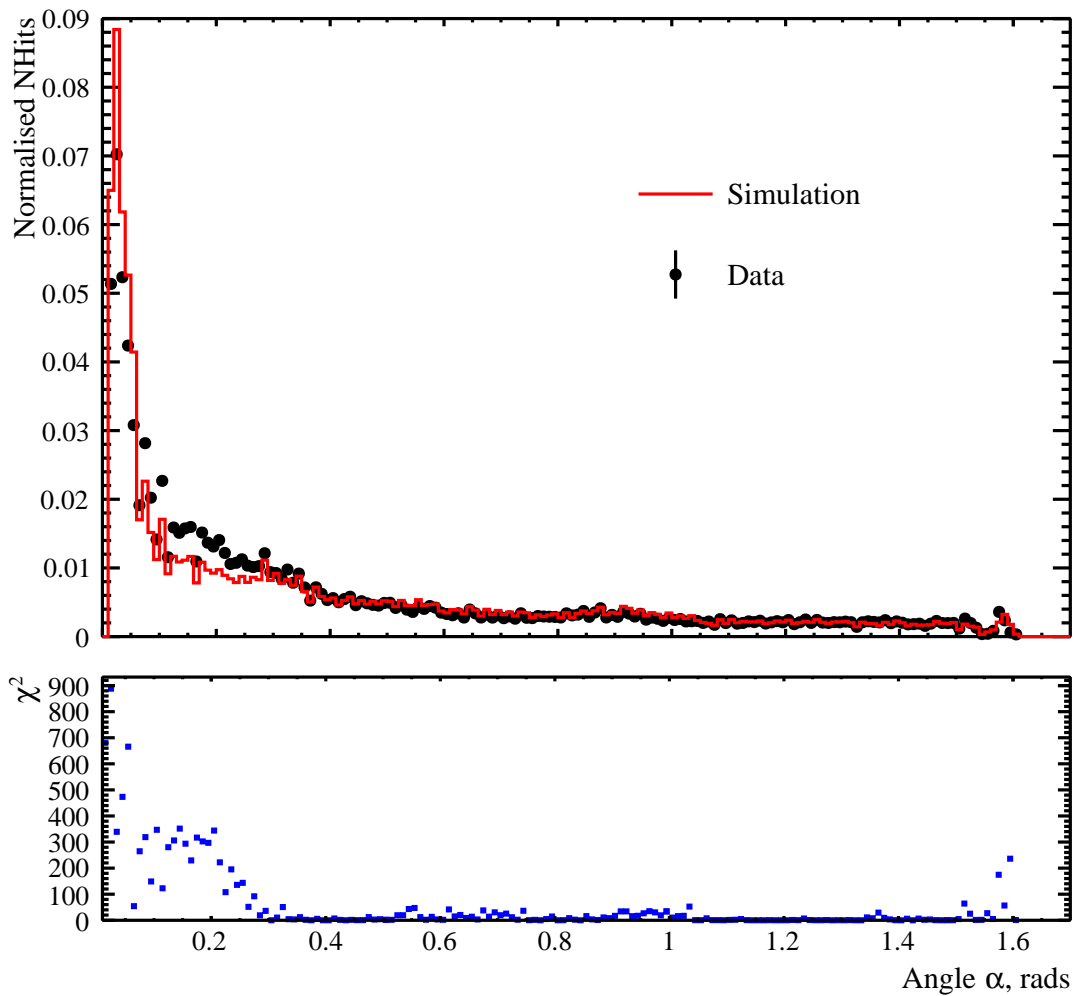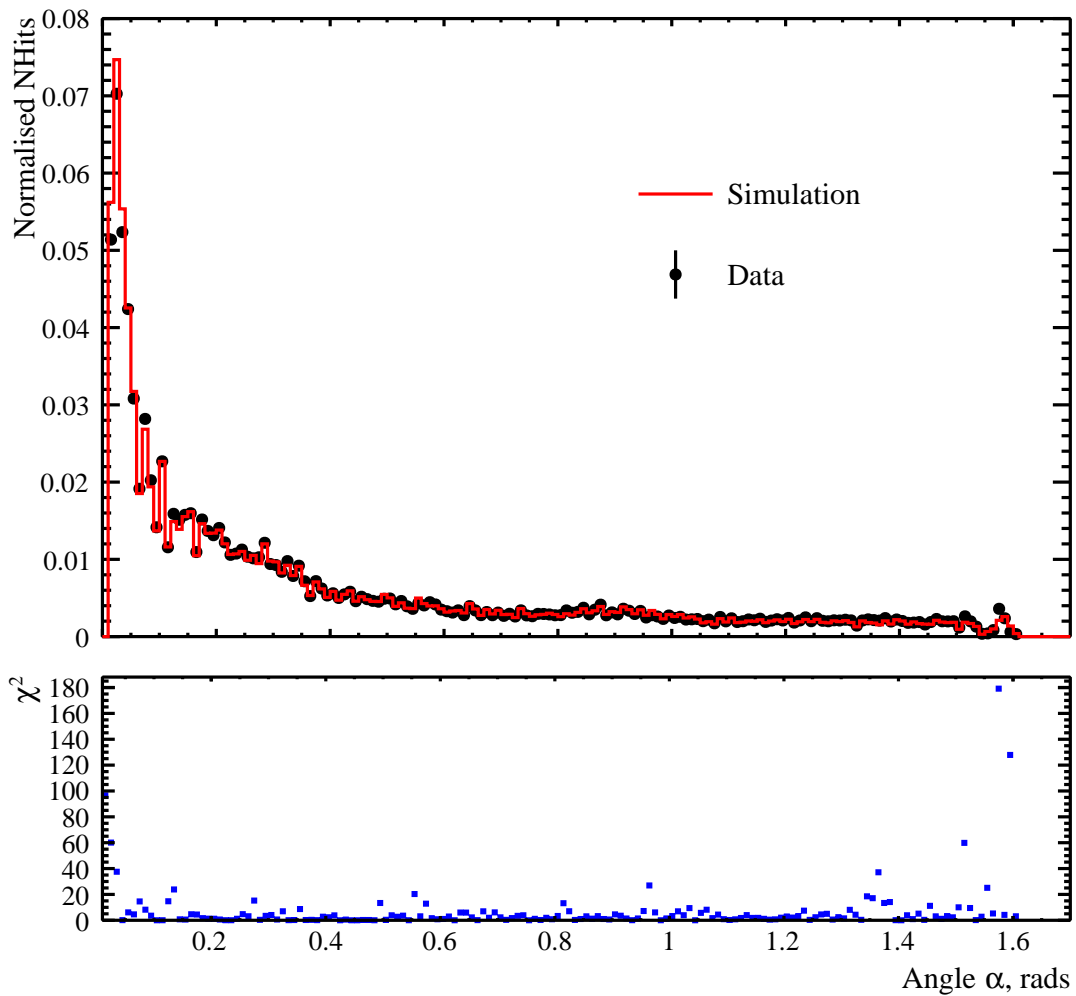
Figure 7.36: Simulation created at 375 nm using the 2D beam profile of FS007 made at 375 nm is shown in red, compared to the matching 375 nm data subrun from the same fibre in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 7.08.
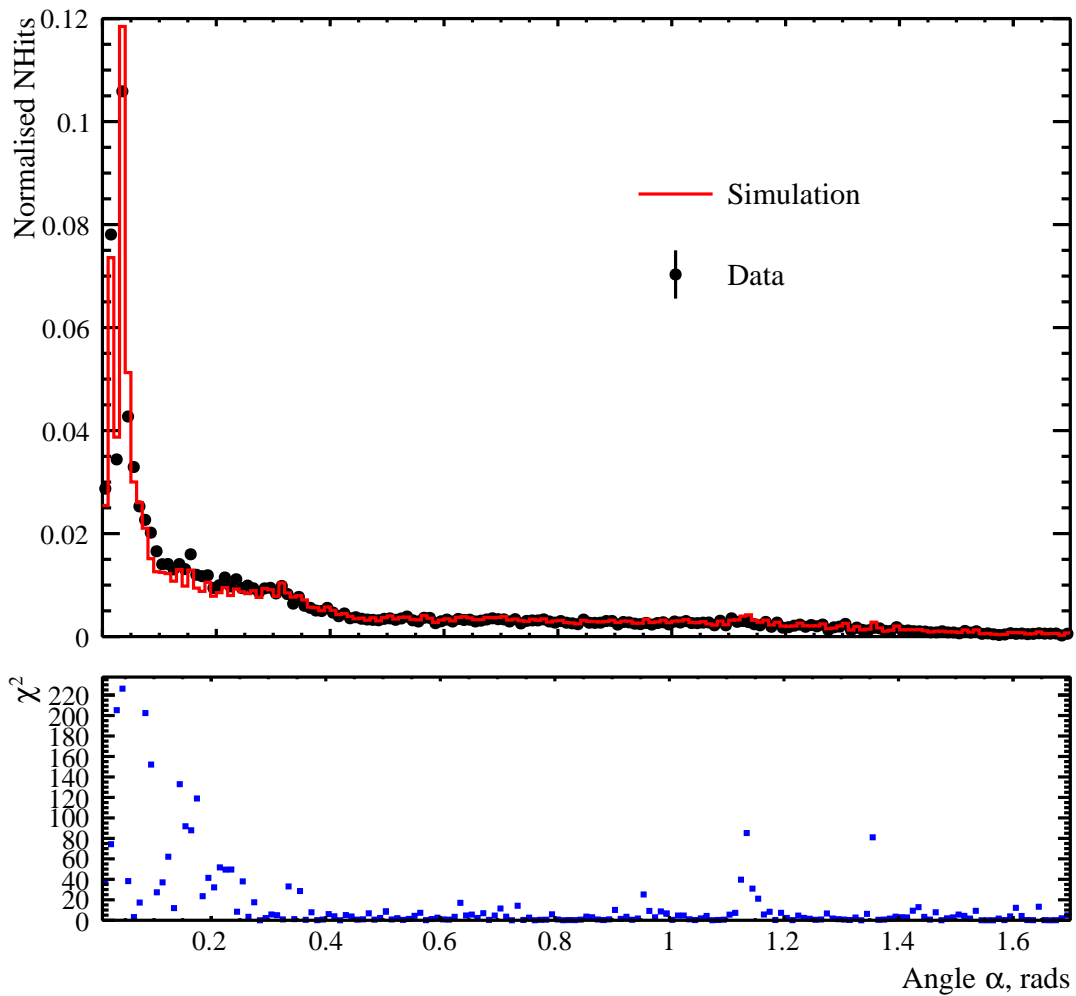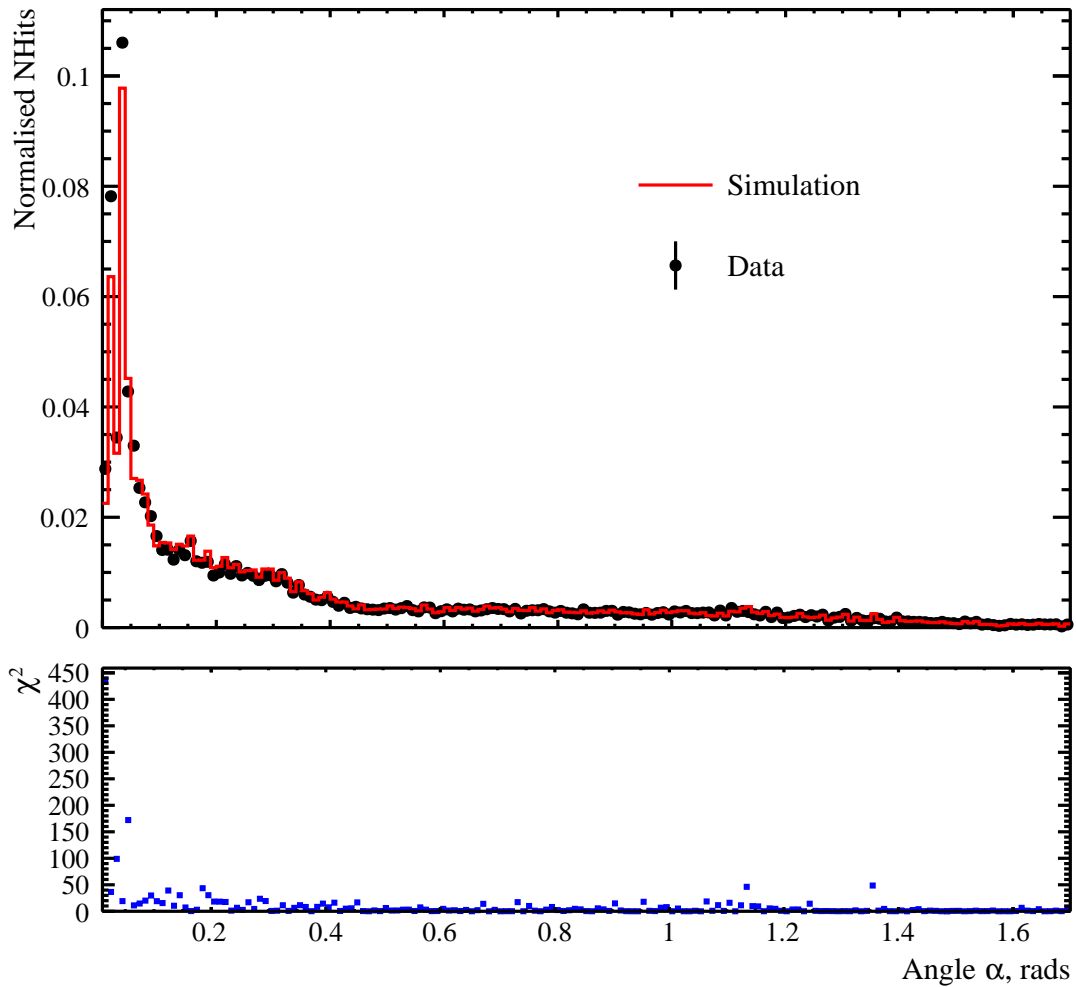
Figure 7.37: Simulation created at 375 nm using the 2D beam profile of FS237 made at 375 nm is shown in red, compared to the matching 375 nm data subrun from the same fibre in black. Poisson error bars are present on the data points, but are too small to see. Both data and simulation have been normalised so that the area under each histogram is equal to one. The χ^2 for each bin has been calculated and is plotted below the main plot, in blue. The overall χ^2_{DOF} is then calculated and, for this plot, is found to be 11.35.

have full coverage of the detector. Therefore, it was decided to treat this mismodelling as a systematic error in the scattering analysis.

Figs. 6.4 - 6.9 show the ratio of p.e. per PMT in data to p.e. per PMT in isotropic simulation. The ropes show up as shadows in both data and simulation. However, in the ratio of data to simulation, the rope in data is a line of shadowed PMTs, but the rope in simulation is a bright line of PMTs (where shadowed translates to a dip in the ratio value and bright a peak). Therefore, the distribution of photons fired by the SMELLIE generator is incorrect. Clearly, this is going to be a more severe effect the closer the rope shadow is to the centre of the beamspot, as this is where the majority of photons are fired.

Evaluating the effect this mismodelling has on the scaling factor analysis is not trivial. While the ropes can be removed in simulation, they, obviously, are immovable in data. Therefore, evaluating this systematic error has been attempted by removing the ropes from the isotropic simulations used in beam profiling and rerunning the analysis as normal, except using these altered beam profiles for the simulations. Obviously, this is only going to remove the rope peaks in the ratio and the dips caused by the ropes in data are going to remain. Therefore, considering the result of this in isolation will not reveal the full effect of this systematic.

Firstly, this process is done for a fibre in which ropes can be seen in the beam profile, but not close to the centre. FS007, shown in Fig. 6.4c has been chosen in this case. The closest ropes to the beamspot centre are approximately 0.3 rads away. Unsurprisingly, the scaling factor found using beam profiles made with no ropes in the isotropic simulations (1.31 ± 0.07) is entirely statistically consistent with the original analysis (1.25 ± 0.07).

Next, the process is repeated for a fibre in which rope shadows pass closer to the beamspot centre. The fibre chosen is FS225 (Fig. 6.6c), in which a rope shadow passes with 0.1 rads of the beamspot centre in both data and simulation, albeit not in the same location. In this case, the scaling factor calculated using the altered beam profile is 0.44 ± 0.07, whereas the original analysis gave a result of 0.58 ± 0.08. This decrease is also compatible with statistics.

However, in order to more fully understand any potential systematic error caused by the mismodelling of the ropes, another fibre has to be considered. For FS225, the rope shadow in simulation is closer to the beamspot centre than in data. It is desirable to create an altered profile and rerun the analysis for a fibre in which the opposite is true.

This is the case for FS207 (Fig 6.5b), in which the rope shadows pass between 0.1 and 0.2 rads away from the beamspot centre. (It is worth noting that both shadows for FS207 are further from the beamspot centre than for FS225 and therefore, any effect on the scaling factor is expected to be smaller as the proportion of photons in the shadows is lower. Ideally a fibre with rope shadows the same distance away from the beamspot centre as FS225 but in the opposite order would have been used, but such a fibre does not exist.) For FS207, the scaling factor calculated using the beam profile derived from simulation without ropes is 1.61 ± 0.15, whereas the original analysis gave a result of 1.38 ± 0.16. This is, like FS225, a change in the scaling factor that is compatible with statistics. Considering that any effect for FS207 is expected to be smaller than for FS225, this is the expected result.

Therefore, the ropes are not considered as a source of systematic errors in this analysis, especially as FS225 has a rope shadow pass extremely close to the beamspot and can consequently be considered as a worst-case scenario. It is possible, that with an increase in statistics, the effect would no longer be negligible. Therefore, understanding the mismodelling of the ropes is still considered as a source of future work, discussed in Chapter 8.

7.4.4 Beam Profiling Region Overlapping Selection Region

So far in the scattering analysis, no consideration has been taken to ensure that the regions used for beam profiling and the selection region for the scattering analysis do not overlap. However, this should be considered: in the beam profiling region, the simulation has been tuned to match data artificially.

Therefore, to evaluate the effect caused by these two regions overlapping, the scattering analysis for all fibres at 435 nm has been repeated. (All fibres can be used as no simulation needs to be run specifically for this test.) Everything is identical to the original analysis except that the selection region has been terminated at 60° (as a lower limit). This is where the beam profiling region extends to in angle. Any hits at angles below 60° in the scattering analysis are now simply counted as part of the total NHits and not as NHits in the selection region.

For all of the fibres, the results from the analysis in which the region is terminated at 60° are completely statistically consistent with the original analysis results, as shown in Table 7.5. Therefore, the conclusion of this study is that the overlap of the beam profiling and selection regions is a completely negligible systematic error, within the resolution available given the statistical errors.

Fibre	Original Scaling Factor	Terminated Region Scaling Factor
FS007	1.23 ± 0.05	1.26 ± 0.07
FS107	1.32 ± 0.08	1.44 ± 0.11
FS207	1.30 ± 0.12	1.43 ± 0.16
FS025	1.20 ± 0.05	1.19 ± 0.07
FS125	1.31 ± 0.07	1.21 ± 0.08
FS225	1.02 ± 0.08	0.96 ± 0.12
FS037	1.17 ± 0.07	1.16 ± 0.09
FS137	1.24 ± 0.05	1.20 ± 0.06
FS237	1.10 ± 0.07	1.08 ± 0.09
FS055	1.16 ± 0.06	1.12 ± 0.07
FS155	1.12 ± 0.10	1.07 ± 0.14
FS255	1.23 ± 0.05	1.18 ± 0.06
FS193	1.03 ± 0.17	1.16 ± 0.21
FS293	1.23 ± 0.12	1.30 ± 0.18

Table 7.5: The difference in scaling factor resulting from terminating the selection region at 60° for all fibres.

7.4.5 Effect of Changing the Selection Region

In a similar vein to the above, the next systematic to investigate is whether there is any systematic error introduced by using a conservative selection region for all fibres, rather than using the individually optimised regions. This conservative region, as the name implies and Fig. 7.13 shows, is always completely contained within the individually optimised regions. Therefore, any error is only associated with moving the two cuts inwards, towards a more restrictive signal selection, and not outwards. As discussed in Section 7.1.2, this is due to the geometry of the detector in the time residual versus α angle space.

In order to evaluate this error, the scaling factors from the 495 nm analyses conducted with both the individually optimised and conservative regions are compared for each fibre. The statistical errors on the two results for each fibre are quadratically combined and then the difference in scaling factors is calculated in units of this combined error. This is referred to as the sigma difference between the two scaling factors. The sigma differences for all the fibres are then used to fill a histogram and a Gaussian is fitted to it. If there is no systematic error introduced by reducing the selection region in this way, a Gaussian with a mean of zero and sigma of one would be expected to fit this distribution, i.e. the difference in scaling factors exceed the combined statistical errors with the frequency expected from a purely statistical effect.
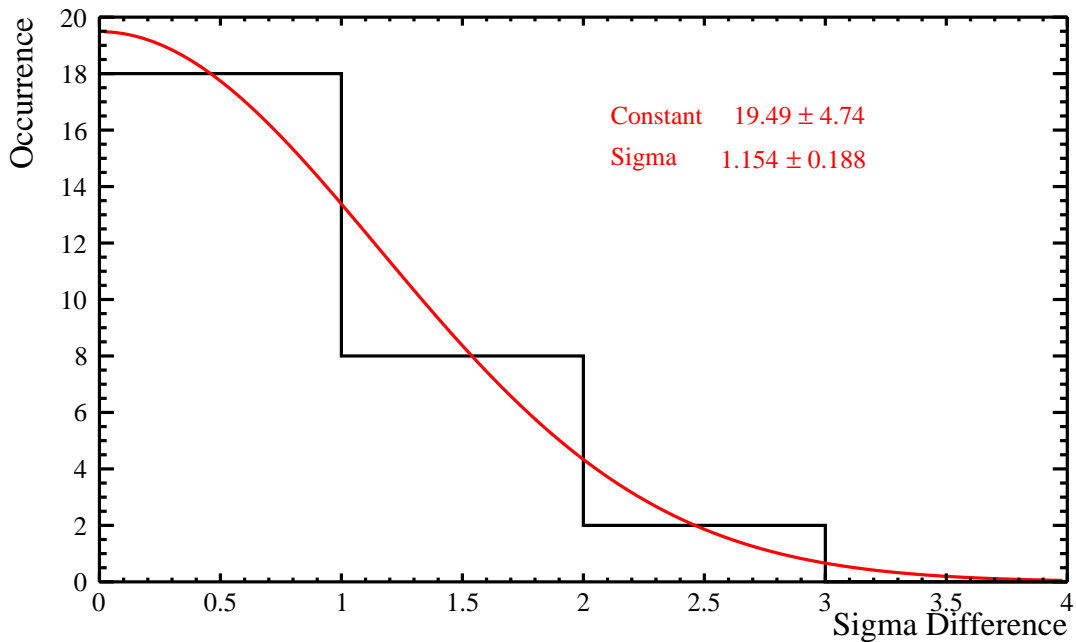
Figure 7.38: Differences between scaling factor results from analyses using the conservative selection region and individually optimised regions for each fibre. This difference is shown in terms of how many multiples of the combined statistical errors the two results are from each other (the sigma difference). A Gaussian in red is fitted to the histogram and the results of this fit are shown in the top right. The mean of the Gaussian has been fixed to zero.

However, as there are only 14 fibres (due to excluding FS093) the simulation datasets have been divided in two in order to increase the number of entries for the Gaussian fits. This was done by taking the even events into one dataset and the odd events into another. The data subruns remain whole. Therefore, there are now 28 entries in the histogram rather than 14, making it easier to fit a Gaussian to the distribution.

This histogram and Gaussian fit are shown in Fig. 7.38. The mean of the Gaussian is fixed to zero while its sigma is allowed to float. The resulting sigma is 1.2 ± 0.2. This is consistent with a sigma of one and therefore it is concluded that moving from an individually optimised region to the conservative region introduces no error other than the increase in statistical error due to the reduction of hits in the selection region.

7.4.6 Shot by Shot Variations in Shot Energy

As discussed in Chapter 5, the CAEN traces can be used to track shot by shot variations in the pulse energy of the lasers and therefore the number of photons

fired into the detector. The height of the CAEN trace (as defined in Section 5.5) is correlated to the p.e. as shown in Fig. 5.32. However, in order to make this plot, by definition, a subrun had to be used in which the laser was not too stable.

For a superK subrun with stable laser behaviour, the equivalent plot is shown in Fig. 5.31. This shows that, as expected, given stable laser running, there is no correlation between the height of the CAEN trace and the cleaned NHits in the detector within a single subrun. The subrun shown in this plot is the data subrun for FS007 at 495 nm used in the scattering analysis.

The CAEN trace height distribution for the same subrun as in Fig. 5.31 is shown in Fig 7.39. The histogram has been fitted with a Gaussian. Naively, it would be expected that the sigma of this Gaussian would be the square root of the mean (i.e. following Poisson statistics). However, the actual sigma of the Gaussian is 32.2 ± 0.3 ADC counts, compared to a mean of 702.4 ± 0.3. The Poisson error on this mean would be 26.5 ADC counts. This is significantly lower (approximately 18%) than the sigma actually observed for this distribution and indicates that the fluctuations observed on the CAEN height are larger than Poisson fluctuations. This is due to laser instability - only a perfect stable laser would cause a purely Poissonian error. Even behaviour that is regarded in this work as stable laser behaviour is not perfectly stable.

In Fig. 7.40, a histogram of cleaned NHits per event is shown for the same subrun. This histogram has also been fitted with a Gaussian. The sigma of the Gaussian is 5.07 ± 0.04 Nhits, compared to a mean of 33.77 ± 0.05. In this case, the fluctuations on the NHits are actually smaller than Poisson fluctuations (by approximately 13%). The level of these fluctuations are compared to a simulation of the same subrun, in which the number of photons fired into the detector has been Poisson fluctuated event by event. In this simulation, even though it is known that the photon number was Poisson fluctuated, the sigma on the Gaussian fitted to the cleaned NHits is smaller than the Poisson error of the mean by 13%. Therefore, it is concluded that in the data subrun shown in Fig. 7.40, the fluctuation in shot energy was actually Poisson distributed.

The narrowness of the cleaned NHits distribution compared to a Poisson distribution is likely due to multi-hits. Ideally, p.e. would be used rather than cleaned NHits. However, the multi-hit correction used in this work calculates an average p.e. and therefore, multi-hit corrected p.e. cannot be used to compare on an event by event basis.
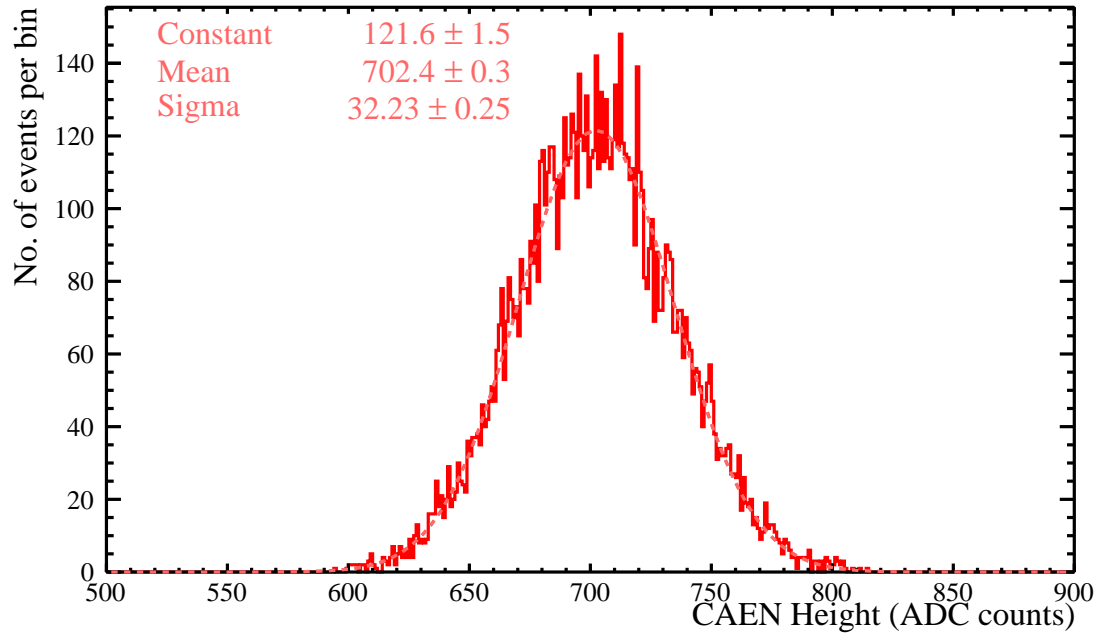
Figure 7.39: The height of the CAEN trace per event for a stable SMELLIE subrun. A Gaussian has been fitted to this distribution and the results of the fit are shown in the top left of the plot.
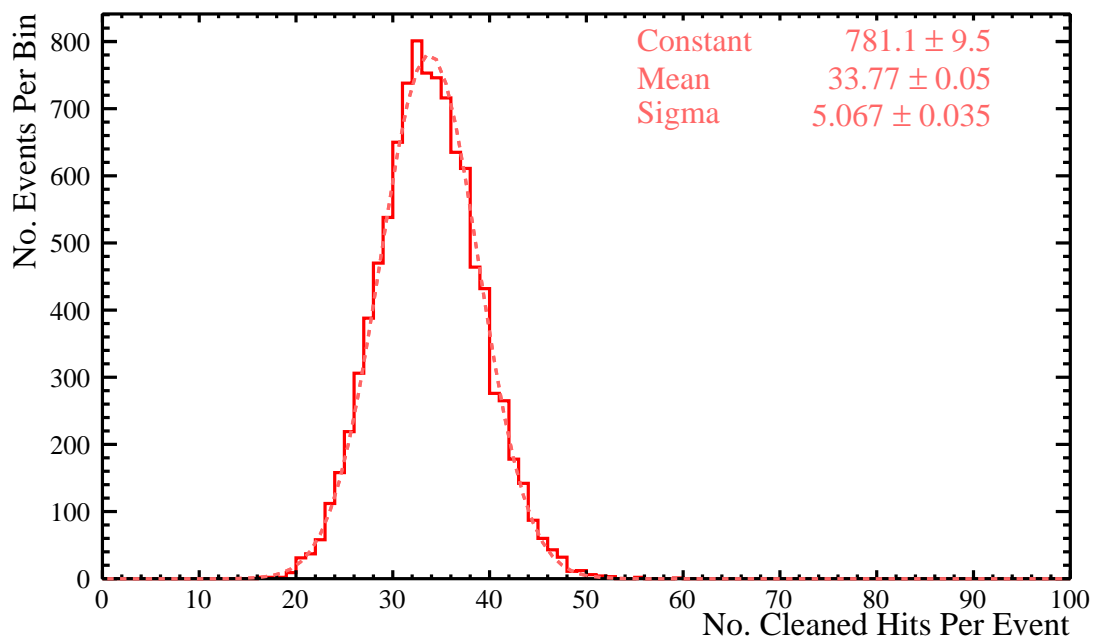

Figure 7.40: The cleaned NHits per event for a stable SMELLIE subrun. A Gaussian has been fitted to this distribution and the results of the fit are shown in the top right of the plot.

As mentioned, in the SMELLIE generator, Poisson fluctuations are used to vary the pulse energy. This matches what is observed in data when the laser is stable. Therefore, for subruns with stable behaviour (i.e. subruns used for the scattering analysis), shot by shot variations in shot energy do not contribute to a systematic error.

For subruns in which the laser does not exhibit stable behaviour, it would be possible to correct the instability of the lasers using the correlation shown in Fig. 5.32. This has not been trialled in this work. However, it has been found that the height of the CAEN traces fluctuate more than would be expected purely from Poisson fluctuations when the laser behaviour is considered 'stable' in this work. Therefore, it would not be possible to correct the shot by shot variations in photons back to the level of Poisson fluctuations using the CAEN traces. On the other hand, the systematic error arising from mismodelling the overall number of photons has been investigated in Section 7.4.1 and it was shown that an error on the level of 42% in the photon number (albeit a shift in the mean photon number rather than a fluctuation) causes no discernible systematic error. This would indicate that the analysis is robust with regards to photon number and that a significant amount of fluctuation could be tolerated. Further testing would be required to demonstrate this more quantitatively.

7.4.7 Mismodelling of the Time Distribution

As seen in Fig. 5.10, despite the data driven definition of the time residuals, the sigma of the prompt peak for the superK laser is not consistent between data and simulation. The sigma in data is approximately 0.35 ns smaller than in simulation. (Whilst this is also true for the PQ375 laser, the difference is significantly smaller, at approximately 0.08 ns.) It is suspected that this additional width arises from the timing distribution of the laser specified in simulation. As shown in Fig. 4.4, the superK laser is modelled as having a timing distribution which is approximately a top hat with a width of 2 ns. However, this was measured by the manufacturers using the full bandwidth of the superK, whereas for SMELLIE only 10 nm bandwidths are used. This will effect the timing distribution as photons of different wavelengths propagate through the optical fibres at different speeds. In addition, photons of different wavelengths will be produced at different times in the super continuum generation, meaning that even without considering propagation through fibres, a shorter pulse would be expected for a narrower bandwidth.

Therefore, to evaluate how much of an error the mismodelling of the superK time distribution introduces to the analysis, the analysis of FS007 and FS237 at 495 nm
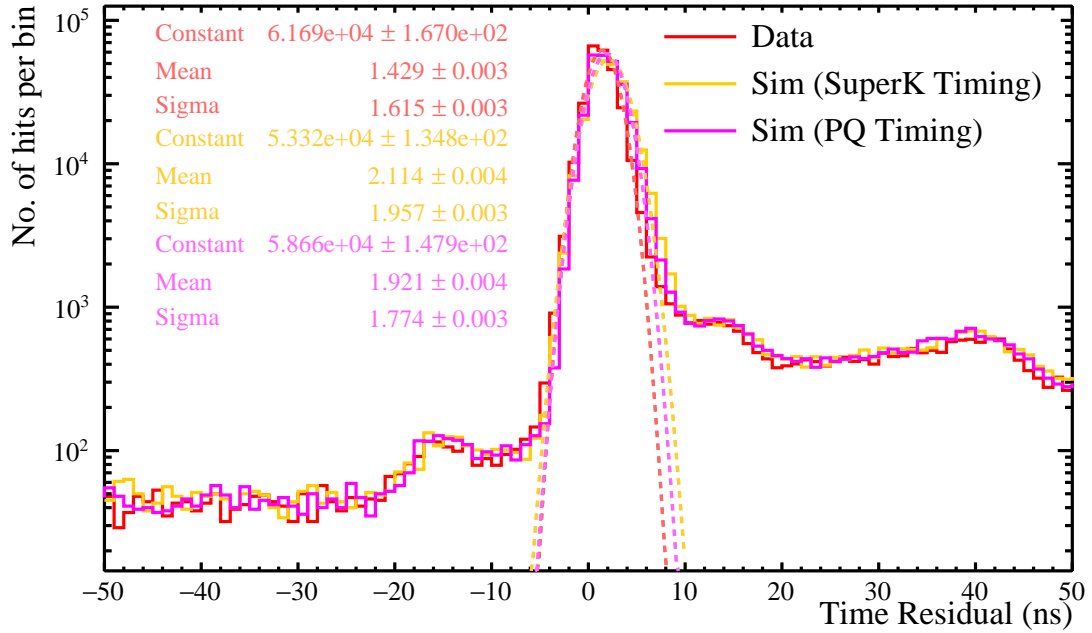
Figure 7.41: Data driven time residuals between data and simulations of the superK laser at 495 nm using FS007. The simulation has been run twice, once using the timing distribution of the superK laser and once using the timing distribution of the PQ495. The prompt peak of all three time residuals are fitted with Gaussians functions (dotted lines) and the results of those fits are shown in the top left of the plot.

(using the superK) has been repeated, but the superK timing distribution is replaced with the PQ495 timing distribution. This distribution as input into simulation is much narrower, with a width of approximately 0.25 ns (see Fig 4.4). Therefore the time residual prompt peak is dominated by the PMT TTS.

The time residuals in data, in simulation with the superK timing distribution and in simulation with the PQ495 timing distribution are shown for FS007 in Fig. 7.41. The prompt peak in the three time residual spectrums in this plot have been fitted with Gaussians. While the sigmas for data and the simulation using the PQ495 timing distribution are not consistent within the errors of the fit, they are over two times closer to each other than data and the simulation using the superK timing. Investigating the systematic error introduced into the scaling factor caused by the difference between the PQ495 and superK timing difference will give a good indication of the systematic introduced by the larger difference between the data and the superK timing spectra.

The result of this analysis for both fibres resulted in scaling factors which were completely statistically consistent with the results from the original analysis (for

FS007, 1.22 ± 0.07 compared to the original 1.25 ± 0.07 and for FS237, 0.99 ± 0.06 compared to 0.97 ± 0.07). This is not a surprise, considering that the conservative selection region was chosen to move the selection region further away from the direct beam, where the prompt peak of the time residual is. Additionally, the level of difference in the time profiles is 0.35 ns, which is smaller than the 1 ns resolution by which the cuts are moved. Therefore, the systematic introduced from this time distribution mismodelling is negligible.

7.4.8 Mismodelling of Fibre Direction

The fibre directions (as discussed in more detail in Chaper 4) were initially set in simulation using the knowledge of the installation location and orientation of the SMELLIE fibre installation plate. However, from data it was observed that this fibre direction did not perfectly match detector data due to freedom for the fibres to move within the plate and for the plate itself to tilt. This is an effect of a couple of degrees; however, over the distance across the detector, this has a noticeable effect on the beamspot locations. Therefore, using detector data, the fibre directions in RAT were tuned (see Section 4.1.2) [105].

It is also the case that one advantage of the 2D beam profiles is that it is not important whether the beamspot centre is located at $(0,0)$ (as defined in polar coordinates for each fibre). Since the profiles are data driven, the entire beam will be located at the same coordinates as it is in data. If the profile was symmetric in azimuthal angle, an offset from $(0,0)$ would seriously distort the beamspot. However, a full 2D profile gives the freedom to not require centred profiles as it is possible for the highest intensity region to be not only at a specified polar angle, but also at a specified azimuthal angle.

On the other hand, the scattering analysis is still only performed in one spatial direction, that of the polar angle. Azimuthal angle is simply summed over. Therefore, a fibre direction which is wrong in the RAT database will create a time residual versus α angle plot in which the beamspot is not at $\alpha = 0°$, but is smeared out in α.

To evaluate the effect this could have on the scaling factor, the fibre directions in RATDB are restored to their original values - those determined purely by the installation node and pointing angle of the fibre, rather than using the direction derived from detector data calculated by T. Tunstall, discussed in Section 4.1.2 and [105]. The analysis is then rerun over the existing simulation. This creates a situation in which the fibre beamspot is a realistic distance away from the $\alpha = 0°$ position on the

angle axis of the time residual versus α angle plot (while avoiding having to rerun simulation).

The fibres chosen to test this effect are FS007 and FS125. The direction vectors for these fibres underwent a typical amount of tuning ($0.72°$ and $2.41°$, respectively [105]). The fibres which were mounted in incorrect positions, such as those on the 37 node, have been ignored in this test, as the adjustment required for these fibres was a matter of position as well as direction. These switched fibres have proved easy to spot from comparisons between detector data and simulation, meaning that such large corrections are not a concern when considering systematic errors to the scaling factor.

For both FS007 and FS125 at 495 nm, the scattering analysis with altered fibre directions resulted in scaling factors of 1.25 ± 0.07 and 1.04 ± 0.06, which are entirely statistically consistent with the original results of 1.25 ± 0.07 and 1.06 ± 0.06 respectively. Therefore, mismodelling of the fibre directions is not considered to contribute a systematic error to this analysis.

7.4.9 Effect of Scattering Model Assumptions

In RAT, as discussed in Chapter 3, scattering is modelled as purely Rayleigh scattering. This determines both the angular and wavelength dependence of the scattering. The angular dependence is modelled through Geant4. However, the wavelength dependence is modelled through the RAT optics database tables.

Despite expecting to observe Rayleigh scattering in the detector, as discussed in Chapter 3, the possibility should be considered that there are other types of scattering happening in the detector. Investigating the angular distribution of the scattering is a future extension to this work and could be used to distinguish scattering types (see Chapter 8). However, due to the wavelength range available to the analysis through the PQ and superK lasers, the easiest way to investigate whether there is a non-Rayleigh scattering component is to investigate the wavelength dependence, as done in Section 7.1.3. As discussed, for Rayleigh scattering, because the RAT model already includes the λ^4 dependence of the scattering length, the scaling factor should be completely independent of wavelength.

Therefore, the aim of this section is to investigate whether the wavelength dependent analysis is a legitimate technique to probe the Rayleigh nature of the scattering in the detector. In order to do this, the model of the scattering wavelength dependence in RAT is changed to a model which approximates Mie scattering (at least in the case where the scatterer size is much larger than the wavelength of the light) - flat
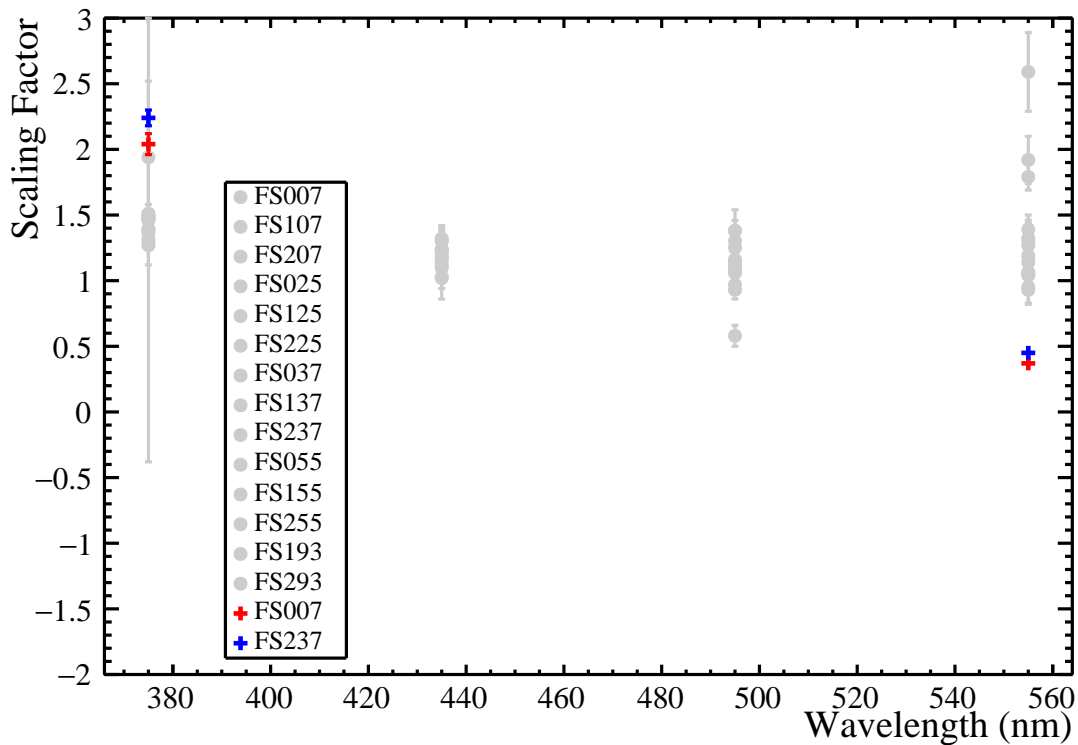
Figure 7.42: Testing the effect of modelling the scattering in RAT as wavelength independent. The original RAT optics have been altered so that the scattering length of internal water for 420 nm applies to all wavelengths and the absorption lengths have been adjusted to keep the attenuation lengths constant. The grey points are the original results of the scattering analysis for each fibre and the two colour crosses (red for FS007 and blue for FS237) show the results of the two fibres at the two wavelengths used to test implementing the change to the optics.

across wavelengths [111]. To be precise, the scattering length of the original internal water at 420 nm has been applied to all wavelengths. Accordingly, the absorption length of internal water has been adjusted in order to keep the original internal water attenuation lengths constant.

The scattering analysis is then repeated at 375 and 555 nm as these are the extremes of wavelength used in the original scattering analysis. As with the majority of the other systematics investigated, FS007 and FS237 are used. The results are shown in Fig. 7.42, overlain on the results from the original scattering analysis (the points in grey).

Even without any quantitive analysis of these results, it is apparent that the hypothesis of a scaling factor that is independent of wavelength is not applicable in this scenario. If the new results are compared to a scaling factor of 1.28 ± 0.05 and the sigma differences are calculated, as defined in Section 7.4.5, the sigma differences

Origin of Systematic	Error (%)
Beam Normalisation	N/A
Beam Profile Mismatch	11
Ropes in Beam Profile	N/A
Region Overlap	N/A
Selection Region	N/A
Time Distribution	N/A
Fibre Direction	N/A
Shot Energy Variation	N/A
Overall	11

Table 7.6: The origin of each systematic error investigated and its percentage effect on the resultant scaling factor.

range from 8 – 13 sigma. However, sigma differences for the same subruns when the original optics are used range from 0 – 3 sigma.

This test is not designed to give a numerical value of a systematic error. Rather, it is done in order to demonstrate that if there were another type of scattering dominant in the internal water in the SNO+ detector, the scattering analysis as performed in this chapter could be sensitive to it. This has been shown and is reinforced by the scattering analysis having discounted a decrease in the scattering length of the acrylic by a factor of 100.

7.4.10 Overall Systematic Error

Systematic errors from various sources within the scaling factor analysis have been investigated above and the results are shown in Table 7.6. It is interesting to note that the only systematic errors of note are those arising from mismodelling in the beam profiles. This indicates that, time permitting, the beam profiles should be returned to as an area of active work. This will be discussed further in Chapter 8.

As shown in Table 7.6, the total systematic error on the scaling factor applied to the optics of both internal and external water in RAT is 11%. This means that the final result of this analysis is a scaling factor of $1.28 \pm 0.05(\text{stat.}) \pm 0.14(\text{sys.})$.

This concludes the first in-situ measurement of scattering lengths in the internal and external waters of the SNO+ detector - a crucial measurement for analyses going forward. The scaling factors have replaced the default values that previously existed in the RAT database and will be used in physics analyses henceforth. It is likely that water phase provided the only opportunity to make a measurement in the external water and this has been capitalised on. In addition, the wavelength dependence

indicates that the assumption that Rayleigh scattering in the water is dominant is correct. The scattering analysis has also been used to inform the optical model of the acrylic.

Chapter 8

Conclusion

8.1 Future Plans

Following the success of this work, there are several avenues of work that should be pursued. The most obvious of these is that SMELLIE, having been used successfully in water phase, must be used to measure the scattering characteristics of the scintillator cocktail during the scintillator phase of the experiment (and after, of course, the Te-loaded scintillator). In addition to this, there are further improvements that can be made to the water phase results, both in terms of extending the analysis as well as reducing systematic errors.

8.1.1 Scintillator Phase

The most crucial next task for SMELLIE is the measurement of the scattering scaling factor for the scintillator cocktail. For unloaded scintillator, there have been previous measurements of the scattering length [112]. However, the scattering length of the Te-loaded cocktail has only been measured on small scales and, therefore, has a higher level of uncertainty associated with it [87]. Performing the scattering analysis in unloaded scintillator is consequently a crucial step as it will introduce complexities which will be present in the analysis of the final cocktail but were not for UPW. These differences arise from the optical differences between LABPPO and UPW.

Firstly, the attenuation length is significantly shorter in LABPPO than in UPW. The absorption and scattering lengths for LABPPO are shown in Figs. 8.1 and 8.2. (For a comparison with UPW, see Figs 3.10 and 8.3 respectively.) While Fig 8.2 shows the scattering length for the unloaded scintillator cocktail, Fig. 8.1 shows the absorption spectrum for LAB and PPO separately. This is because, for the SMELLIE wavelength range, the absorption length of PPO is longer than for LAB - the LAB
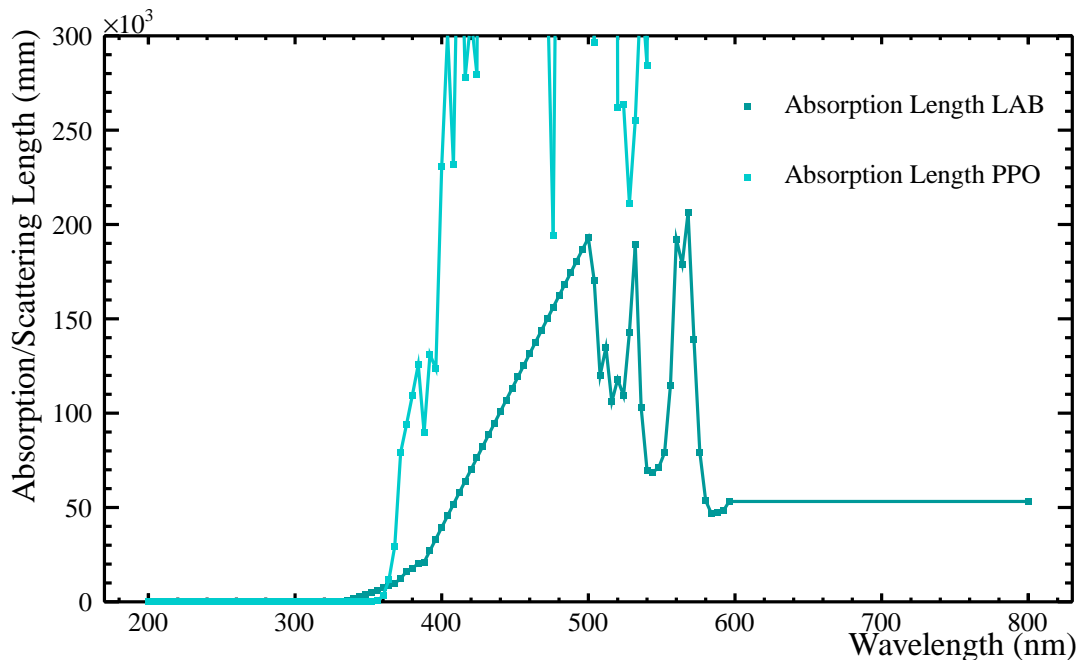
Figure 8.1: Absorption spectra of LAB and PPO as in the optics tables in RATDB. The PPO spectrum is not fully displayed as the LAB spectrum is of particular interest here.

absorption is dominant [87]. For both absorption and scattering lengths, values above 600 nm have been set to a constant value as this is not a wavelength range of interest.

Although absorption can be assumed to occur through LAB, any re-emission from LAB is assumed to be reabsorbed. Therefore, to consider re-emission (which is present in scintillator and absent in water), non-radiative transfer to PPO from LAB must be considered, followed by re-emission from PPO. Re-emission from LAB itself is therefore not treated in RAT. Both the non-radiative transfer to PPO and the quantum yield (likelihood of emission) of PPO must be considered to calculate the overall likelihood of re-emission after a photon has been absorbed. This results in an overall re-emission probability of approximately 0.6 [87].

Before the scattering analysis can be performed, an understanding of how to adjust parameters of the SMELLIE system to result in the desired observed NHits/QHS for any fibre/wavelength/laser combination in scintillator SMELLIE data must be gained. As the absorption and scattering lengths in scintillator are significantly shorter than in water, the shot energy settings must be adjusted to achieve the desired shot energy regions for each fibre/wavelength combination.

The main additional difficulty that is present when performing the SMELLIE analysis in scintillator rather than water is re-emission. This behaviour, in terms of
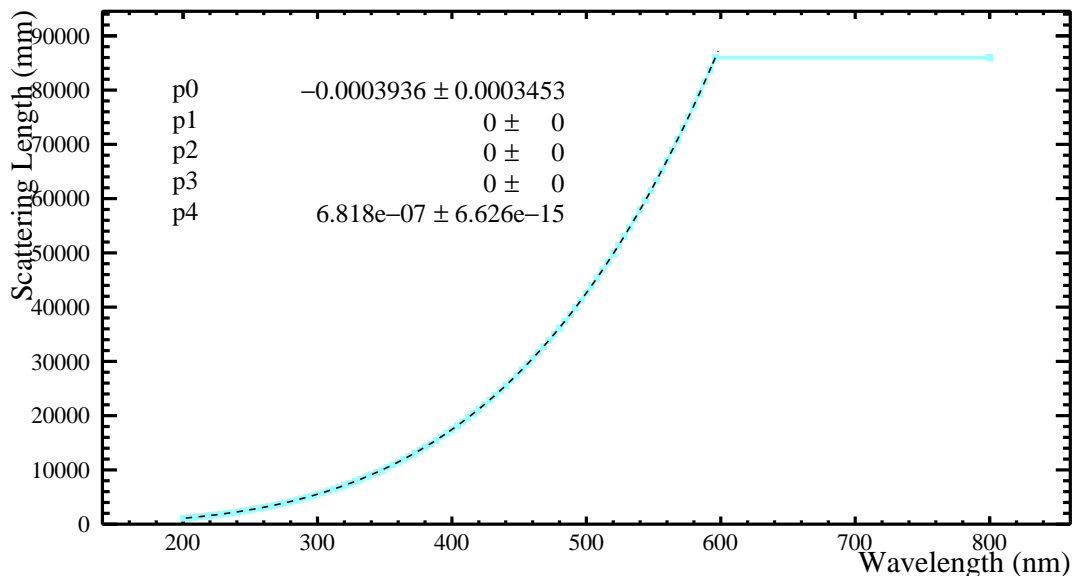
Figure 8.2: Scattering length of LABPPO of a function of wavelength as in the optics tables in RATDB. The spectrum is fitted with a function of the form $y = p0 + p1\lambda + p2\lambda^2 + p3\lambda^3 + p4\lambda^4$ where p1, p2 and p3 are fixed to zero to match the expected Rayleigh scattering wavelength dependence.

detector observables, is very difficult to separate from scattering; the direction of photons originating from SMELLIE is altered by both re-emission and scattering in the detector medium. However, re-emission is isotropic; the photon direction after re-emission has no correlation to the direction prior to re-emission. Therefore, a way to differentiate re-emission and scattering in detector data is to examine the angular distribution of the PMT hits with respect to the initial direction of the photon. However, doing this requires the reconstruction of the scattering point of the photon. There has been initial investigation into this, first by K. Majumdar [53], followed up by C. Ransom [113].

It would also be possible to investigate the angular dependence of the scattering present in the water phase, using data taken during this work (specifically, the high shot energy subruns). This would allow for the development and testing of a method on detector data without the complication of re-emission and would also be able to confirm the conclusion of this thesis that there is no indication of non-Rayleigh scattering present in the UPW.

8.1.2 Time Variation in Water Phase

As well as the scintillator scattering analysis, there remains some further exploration that can be done with the water phase data already taken. One avenue to investigate

is whether there is any time variation in the Rayleigh scattering scaling factor. SMEL-LIE data was taken throughout the water phase, though, as discussed in Section 5.7, not all of the data is suitable for use in the scattering analysis. However, there is data available in the data range from November 2017 to June 2018 that would be suitable for use, though not for all fibre/wavelength/shot energy combinations that were presented in this work. (The analysis in this work uses the last set of data from this range.) Comparing data over a period of time does mean that the detector states are not necessarily identical. However, all of this SMELLIE data was taken in between 'gold' runs for water physics analyses, meaning that as constant a detector state as possible will have been kept. An additional concern is that the beam profiles used in this work were created from data taken at the same time as the data used in the scattering length analysis, ensuring that no systematic difference in detector state was present between the two data sets.

8.1.3   Beam Profiling

As shown in Section 7.4, the largest systematic errors in the Rayleigh scattering scaling factor analysis arise from the beam profiles. Therefore, some focus should be placed on improving the profiles moving forwards, enabling more precise measurements of the scaling factor.

This effort should be focused in two areas. One is to understand whether the mismodelling of ropes in beam profiles is due to incorrect positions of SMELLIE injection points in simulation or due incorrect rope positions. This mismodelling, whether it arises from the rope positions or the fibre positions, is not a discernible source of systematic error in the scattering analysis. However, it is concerning. In this analysis, for reasons discussed in Section 6.2.2, the assumption has been made that it is the rope positions that are incorrect. However, although this can not be confirmed with SMELLIE data, it would be possible using TELLIE data as TELLIE beamspots have entire detector coverage. If the rope positions are systematically wrong throughout the detector, these should be corrected in simulation. Conversely, if it becomes apparent that the fibre positions are located incorrectly in simulation, these should be corrected.

More importantly, the other area that should be focused on is the mismodelling of the profile itself. In Section 7.4.2, it was shown that simulations created using the beam profile at the same wavelength as the profile was created at caused negligible systematic errors in the scattering analysis. However, creating simulation at a different wavelength to the one used to create the profiles lead to a systematic error of 11%

in the scaling factor. Therefore, to minimise this effect, one option would be to create a range of beam profiles at the wavelengths usually used in analysis. This would resolve mismodelling arising from wavelength dependent light propagation through the fibre and collimator. Alternatively, beam profiles could be created at as high a wavelength as possible, to minimise the amount of scattered light included in the beam profile, as discussed in Section 7.4.2.

Water data is required for beam profiling as the increased scattering and addition of re-emission in scintillator data would make the profiles more difficult to determine, effectively blurring them. Therefore, data taken during the course of this work would have to be used.

8.1.4 Continued Commissioning

As discussed in Chapter 5, there remain some commissioning issues that were not completely satisfactorily resolved during the water phase. Crucially, the trigger jitter due to latching to the 100 MHz clock needs to be fixed before SMELLIE data is taken in scintillator phase. For water, a data driven t_0 for each event was used. However, this analysis correction would be more difficult in scintillator, in which the attenuation length is significantly shorter than water. Hence, at some wavelengths, it would not be possible to reliably use a direct photon hit to derive t_0. A hardware fix is therefore preferred - the chip in TUBii that provides the 'asynchronous' delay will be replaced with one that is truly asynchronous before scintillator phase data taking commences.

8.2 Summary

In this work, a calibration of scattering behaviour in the SNO+ detector medium in water phase was conducted. This involved measuring the Rayleigh scattering length of the internal UPW, via a measurement of the divisive scaling factor applied to the default values in RAT (which were extrapolated from SNO values). The scaling factor was found to be 1.28 ± 0.05(stat.)± 0.14(sys.). The effect of this measurement of the scaling factor on the UPW scattering length is shown in Fig. 8.3.

This measurement was made using 14 fibres across the detector in order to utilise various different path lengths through the detector materials. The resultant scaling factor was then combined with the attenuation length measured by the laserball analysis in order to determine the scaling factor for the absorption length of internal UPW. These two scaling factors were also applied to external water. By measuring the Rayleigh scattering scaling factor at a variety of wavelengths throughout the
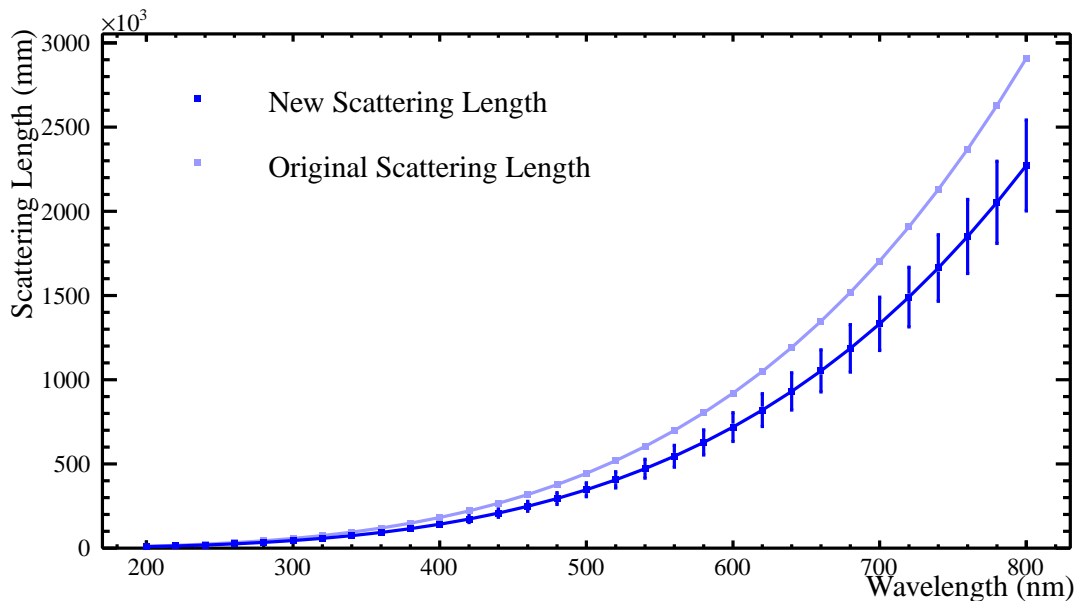
Figure 8.3: The UPW scattering length as in RAT before and after the measurement made in this work. The full range of the RAT scattering length is shown; however, the measurement in this work was only in the wavelength range of 375 - 555 nm.

optical spectrum (375, 435, 495 and 555 nm), the wavelength dependence of the scattering length length was investigated. This was shown to be consistent with Rayleigh scattering behaviour, leading to the conclusion that Rayleigh scattering is the only significant source of scattering in the SNO+ UPW.

As well as characterise the scattering behaviour of UPW, the SMELLIE scattering analysis was used to inform the choice of optical constants for the acrylic after the external laserball runs were analysed.

In order to achieve any of these results, SMELLIE was also commissioned during this work. Before 2015 it had only been fired into a partially filled detector a handful of times and crucial components of the system, such as the superK laser, had yet to be installed. Understanding the system behaviour and making necessary hardware adjustments was therefore a crucial aspect to this work. Additionally, measurements have been made of the beam profiles using detector data, after discovering that the 1D profiles previously measured were not an accurate enough representation of the beams seen in the detector to conduct the scattering analysis.

In conclusion, this work has resulted in characterisation of the scattering present and a wavelength dependent measurement of the scattering length of the UPW detection medium in the first phase of the SNO+ experiment. Also, essential commissioning of the system needed for both that measurement and for use in future phases has been

conducted. Therefore, building upon the understanding developed in this work will enable equivalent measurements in the future two phases, in less well understood and more optically complex mediums.

References

[1] Wolfgang Pauli, "Pauli letter collection: letter to Lise Meitner", Typed copy (1930), URL https://cds.cern.ch/record/83282

[2] J. Chadwick, "The intensity distribution in the magnetic spectrum of beta particles from radium (B + C)", Verh. Phys. Gesell. vol. 16 pp. 383–391 (1914)

[3] E. Fermi, "An attempt of a theory of beta radiation. 1.", Z. Phys. vol. 88 pp. 161–177 (1934)

[4] H. Bethe and R. Peierls, "The neutrino", Nature vol. 133 (3362) p. 532 (1934)

[5] F. Reines and C. L. Cowan, "The Neutrino", Nature vol. 178 (4531) pp. 446–449 (1956)

[6] C. L. Cowan et al., "Detection of the Free Neutrino: a Confirmation", Science vol. 124 (3212) pp. 103–104 (1956)

[7] G. Danby et al., "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos", Phys. Rev. Lett. vol. 9 pp. 36–44 (1962)

[8] K. Kodama et al. (DONUT), "Observation of tau neutrino interactions", Phys. Lett. vol. B504 pp. 218–224 (2001)

[9] S. Schael et al. (ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group), "Precision electroweak measurements on the Z resonance", Phys. Rept. vol. 427 pp. 257–454 (2006)

[10] T. D. Lee and C. N. Yang, "Question of parity conservation in weak interactions", Physical Review vol. 104 (1) pp. 254–258 (1956)

[11] C. S. Wu et al., "Experimental test of parity conservation in beta decay [5]", Physical Review vol. 105 (4) pp. 1413–1415 (1957)

[12] R. P. Feynman and M. Gell-Mann, "Theory of the fermi interaction", Physical Review vol. 109 (1) pp. 193–198 (1958)

[13] E. C. G. Sudarshan and R. E. Marshak, "Chirality invariance and the universal Fermi interaction [6]", vol. 109 (5) pp. 1860–1862 (1958)

[14] J. J. Sakurai, "Mass reversal and weak interactions", Il Nuovo Cimento vol. 7 (5) pp. 649–660 (1958)

[15] M. Goldhaber, L. Grodzins, and A. W. Sunyar, "Helicity of Neutrinos", Phys. Rev. vol. 109 pp. 1015–1017 (1958)

[16] L. Landau, "On the conservation laws for weak interactions", Nuclear Physics vol. 3 (1) pp. 127–131 (1957)

[17] T. D. Lee and C. N. Yang, "Parity Nonconservation and a Two-Component Theory of the Neutrino", Physical Review vol. 105 (5) pp. 1671–1675 (1957)

[18] A. Salam, "On parity conservation and neutrino mass", Il Nuovo Cimento vol. 5 (1) pp. 299–301 (1957)

[19] J. N. Bahcall, "Neutrinos from the Sun", Scientific American vol. 221 (1) pp. 28–37 (1969)

[20] J. G. Learned, "The Atmospheric Neutrino Anomaly: Muon Neutrino Disappearance", in "Current Aspects of Neutrino Physics", pp. 89–130, Springer Berlin Heidelberg (2000)

[21] M. Tanabashi and others (Particle Data Group), "Review of Particle Physics", Phys. Rev. D vol. 98 p. 030001 (2018)

[22] Y. Fukuda et al., "Evidence for oscillation of atmospheric neutrinos", Physical Review Letters vol. 81 (8) pp. 1562–1567 (1998)

[23] B. Aharmim et al., "Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory", Physical Review C - Nuclear Physics vol. 88 (2) p. 025501 (2013)

[24] A. Y. Smirnov, "The MSW Effect and Matter Effects in Neutrino Oscillations", Physica Scripta vol. T121 pp. 57–64 (2005)

[25] P. F. de Salas et al., "2020 global reassessment of the neutrino oscillation picture", JHEP vol. 02 p. 071 (2021)

[26] K. Abe et al., "Constraint on the matterantimatter symmetry-violating phase in neutrino oscillations", Nature vol. 580 (7803) pp. 339–344 (2020)

[27] S. Adrian-Martinez et al. (KM3Net), "Letter of intent for KM3NeT 2.0", J. Phys. vol. G43 (8) p. 084001 (2016)

[28] M. Fertl, "Review of absolute neutrino mass measurements", Hyperfine Interactions vol. 239 (1) p. 52 (2018)

[29] V. N. Aseev et al. (Troitsk), "An upper limit on electron antineutrino mass from Troitsk experiment", Phys. Rev. vol. D84 p. 112003 (2011)

[30] S. Dell'Oro and S. Marcocci, "Recent results from cosmology and neutrinoless double beta decay", Journal of Physics: Conference Series vol. 718 p. 062012 (2016)

[31] E. Majorana, "Teoria simmetrica dell'elettrone e del positrone", Il Nuovo Cimento vol. 14 (4) pp. 171–184 (1937)

[32] Stefano Dell'Oro, Simone Marcocci, Matteo Viel, and Francesco Vissani, "Neutrinoless double beta decay: 2015 review", Adv. High Energy Phys. vol. 2016 p. 2162659 (2016)

[33] S. M. Bilenky, J. Hosek, and S. T. Petcov, "On Oscillations of Neutrinos with Dirac and Majorana Masses", Phys. Lett. vol. 94B pp. 495–498 (1980)

[34] S. Weinberg, "Baryon- and lepton-nonconserving processes", Physical Review Letters vol. 43 (21) pp. 1566–1570 (1979)

[35] M. Goeppert-Mayer, "Double Beta-Disintegration", Phys. Rev. vol. 48 pp. 512–516 (1935)

[36] L. Maiani, "Selected Topics in Majorana Neutrino Physics", Riv. Nuovo Cim. vol. 37 (08) pp. 417–466 (2014)

[37] V. I. Tretyak and Y. G. Zdesenko, "Tables of double beta decay data - an update", Atomic Data and Nuclear Data Tables vol. 80 (1) pp. 83–116 (2002)

[38] W. H. Furry, "On Transition Probabilities in Double Beta-Disintegration", Phys. Rev. vol. 56 pp. 1184–1193 (1939)

[39] C. Patrignani et al. (Particle Data Group), "Review of Particle Physics", Chin. Phys. vol. C40 (10) p. 100001 (2016)

[40] O. Cremonesi, "Neutrinoless double beta decay: Present and future", Nucl. Phys. Proc. Suppl. vol. 118 pp. 287–296 (2003)

[41] S. Dell'Oro, S. Marcocci, and F. Vissani, "New expectations and uncertainties on neutrinoless double beta decay", Phys. Rev. vol. D90 (3) p. 033005 (2014)

[42] J. J. Gomez-Cadenas and J. Martin-Albo, "Phenomenology of neutrinoless double beta decay", PoS vol. GSSI14 p. 004 (2015)

[43] B. Schwingenheuer, "Status and prospects of searches for neutrinoless double beta decay", Annalen der Physik vol. 525 (4) pp. 269–280 (2013)

[44] A. Gando et al. (KamLAND-Zen Collaboration), "Search for Majorana Neutrinos Near the Inverted Mass Hierarchy Region with KamLAND-Zen", Phys. Rev. Lett. vol. 117 p. 082503 (2016)

[45] M. Agostini et al. (GERDA), "Improved Limit on Neutrinoless Double-β Decay of ^{76}Ge from GERDA Phase II", Phys. Rev. Lett. vol. 120 (13) p. 132503 (2018)

[46] C. Alduino et al. (CUORE), "First Results from CUORE: A Search for Lepton Number Violation via $0\nu\beta\beta$ Decay of ^{130}Te", Phys. Rev. Lett. vol. 120 (13) p. 132501 (2018)

[47] S. Biller and S. Manecki, "Te in Liquid Scintillator for Neutrinoless Double Beta Decay Experiments Te in Liquid Scintillator for Neutrinoless Double Beta Decay Experiments for the SNO+ collaboration", J. Phys.: Conf. Series vol. 888 (2017)

[48] S. Andringa et al. (The SNO+ Collaboration), "Current Status and Future Prospects of the SNO+ Experiment", Adv. High Energy Phys. vol. 2016 pp. 1–21 (2015)

[49] J. Boger et al. (The SNO Collaboration), "The Sudbury Neutrino Observatory", Nucl. Instrum. Methods vol. 449 (1) pp. 172–207 (1999)

[50] K. Singh, "PSUP camera photos", SNO+ DocDB 4331-v1 (2017)

[51] M. Anderson et al. (The SNO+ Collaboration), "Search for invisible modes of nucleon decay in water with the SNO+ detector", Phys. Rev. D vol. 99 p. 032008 (2019)

[52] M. Anderson et al. (The SNO+ Collaboration), "Measurement of the ^{8}B solar neutrino flux in SNO+ with very low backgrounds", Phys. Rev. D vol. 99 p. 012012 (2019)

[53] K. Majumdar, "On the measurement of optical scattering and studies of background rejection in the SNO+ detector", Ph.D. thesis, University of Oxford (2015)

[54] E. Arushanova and A. R. Back, "Physics capabilities of the SNO+ experiment for the SNO+ collaboration", J. Phys.: Conf. Series vol. 888 (2017)

[55] J. Dunger, "Topological and Time Based Event Classification for Neutrinoless Double Beta Decay in Liquid Scintillator", Ph.D. thesis, University of Oxford (2018)

[56] J. Paton (The SNO+ Collaboration), "Neutrinoless Double Beta Decay in the SNO+ Experiment", Prospects in Neutrino Physics (2018): London, UK, December 19-21, 2018 (2019)

[57] J. V. Jelley, "Cerenkov Radiation: its Origin, Properties and Applications", Contemporary Phys. vol. 3:1 pp. 45–57 (1961)

[58] D. F. Alferov, Y. A. Bashmakov, and P. A. Cherenkov, "Radiation from relativistic electrons in a magnetic undulator", Phys. Usp. vol. 32 (3) pp. 200–227 (1989)

[59] I. Coulter, "Modelling and Reconstruction of Events in SNO+ Related to Future Searches for Lepton and Baryon Number Violation", Ph.D. thesis, University of Oxford (2013)

[60] M. F. L'Annunziata, "Handbook of radioactivity analysis [electronic resource]", Elsevier Science, San Diego, 3rd edn. (2012)

[61] S. N. Ahmed, "Physics and engineering of radiation detection [electronic resource]", Elsevier Science, Amsterdam, Netherlands, 2nd edn. (2015)

[62] M. D. Lay and M. J. Lyon, "An experimental and Monte Carlo investigation of the R1408 8-inch photomultiplier tube and associated concentrator to be used in the Sudbury Neutrino Observatory", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 383 (2) pp. 495 – 505 (1996)

[63] Hamamatsu Photonics K.K., "Photomultipler Tubes: Basics and Applications [electronic resource]", Hamamatsu Photonics K.K., 3rd (3a) edn. (2007)

[64] S. D. Biller et al., "Measurements of photomultiplier single photon counting efficiency for the Sudbury Neutrino Observatory", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 432 (2) pp. 364 – 373 (1999)

[65] T. Kroupova, "PMT Transit Time Tuning Note", SNO+ DocDB 5353-v1 (2018)

[66] V. Albanese et al. (The SNO+ Collaboration), "The SNO+ experiment", JINST vol. 16 (08) p. P08059 (2021)

[67] CAEN, "V1720 - 8 Channel 12bit 250 MS/s Digitizer", URL `https://www.caen.it/products/v1720/`, last checked 2021

[68] T. Kaptanoglu, "Reactor Antineutrinos in the SNO+ Water Phase and Detector R and D for Large-Scale Neutrino Detectors", Ph.D. thesis, University of Pennsylvania (2020)

[69] R. Stainforth, "Characterising the Optical Response of the SNO+ detector", Ph.D. thesis, University of Liverpool (2016)

[70] J. Maneira et al., "Calibration of the SNO+ experiment", J. Phys.: Conf. Series vol. 888 (2017)

[71] B. A. Moffat et al., "Optical calibration hardware for the Sudbury Neutrino Observatory", Nucl. Instrum. Methods vol. 554 pp. 255–265 (2005)

[72] J. Cameron, "The Photomultiplier Tube Calibration of the Sudbury Neutrino Observatory", Ph.D. thesis, University of Oxford (2001)

[73] R. Alves et al. (The SNO+ Collaboration), "The calibration system for the photomultiplier array of the SNO+ experiment", JINST vol. 10 (03) p. P03002 (2015)

[74] F. Descamps, "PCA Calibration", SNO+ DocDB 1987-v5 (2016)

[75] A. S. Inacio, "July 2018 Laserball Scan Runlist", SNO+ DocDB 5127-v2 (2018)

[76] A. S. Inacio, "Optical Calibration of the SNO+ Detector with the Laserball in the Water Phase", SNO+ DocDB 5862-v4 (2020)

[77] E. Falk, J. Lidgard, M. I. Stringer, and E. Turner, "Commissioning of ELLIE for SNO+", Prospects in Neutrino Physics: London, UK, December 12-14, 2016 (2017)

[78] A. Falk, "AMELLIE update", SNO+ DocDB 4905-v1 (2018)

[79] S. Steibert, "RAT (is an Analysis Tool) User's Guide", URL `https://rat.readthedocs.io/en/latest/index.html`, last checked 2020

[80] S. Agostinelli et al., "Geant4a simulation toolkit", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol. 506 (3) pp. 250 – 303 (2003)

[81] G. Horton-Smith, "GLG4sim page", URL `http://neutrino.phys.ksu.edu/~GLG4sim/`, last checked 2020

[82] R. Stainforth, "LightPath update (ELLIE Reflections)", SNO+ DocDB 2484-v1 (2014)

[83] H. C. van de Hulst, "Light scattering by small particles", Structure of matter series, Wiley, New York (1957)

[84] C. V. Raman and K. S. Krishnan, "A New Type of Secondary Radiation", Nature vol. 121 (3048) pp. 501–502 (1928)

[85] G. Mie, "Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen", Annalen der Physik vol. 330 (3) pp. 377–445 (1908)

[86] F. R. S. Lord Rayleigh, "XXXIV. On the transmission of light through an atmosphere containing small particles in suspension, and on the origin of the blue of the sky", The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science vol. 47 (287) pp. 375–384 (1899)

[87] M. R. Anderson et al. (The SNO+ Collaboration), "Development, characterisation, and deployment of the SNO+ liquid scintillator", JINST vol. 16 (05) p. P05009 (2021)

[88] C. F. Bohren and D. R. Huffman, "Absorption and scattering of light by small particles [electronic resource]", Wiley, New York (1983)

[89] W. Hergert and T. Wriedt, "The Mie theory : basics and applications [electronic resource]", Springer series in optical sciences ; 169, Springer, Berlin ; London (2012)

[90] E. B. Shybanov, "Some experimental results of the light scattering in water in a new theoretical approach", in Oleg A. Romanovskii and Gennadii G. Matvienko (Eds.), "25th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics", vol. 11208, pp. 1133 – 1137, International Society for Optics and Photonics, SPIE (2019)

[91] M. R. Anderson et al. (The SNO+ Collaboration), "Optical calibration of the SNO+ detector in the water phase with deployed sources", JINST vol. 16 (10) p. P10021 (2021)

[92] A. S. Inacio, "Analysis of the June/July 2018 Internal and External Laserball Scans", SNO+ DocDB 5625-v1 (2019)

[93] B. Moffat, "The Optical Calibration of the Sudbury Neutrino Observatory", Ph.D. thesis, Queen's University (2001)

[94] K. Clark et al., "The SMELLIE Hardware Manual", SNO+ DocDB 3511-v2 (2016)

[95] T. Major, "Measurements of scattering in TeLS", SNO+ DocDB 2081-v1 (2013)

[96] T. Kaptanoglu, "Absorption and scattering of the TeDiol", SNO+ DocDB 3923-v1 (2016)

[97] H. Buiteveld, J. H. M. Hakvoort, and M. Donze, "Optical properties of pure water", in Jules S. Jaffe (Ed.), "Ocean Optics XII", vol. 2258, pp. 174 – 183, International Society for Optics and Photonics, SPIE (1994)

[98] S. Langrock et al., "Scattering Length Monitoring at the SNO+ Detector", SNO+ DocDB 3748-v7: Neutrino 2016 Poster (2016)

[99] PicoQuant, "Picosecond Pulsed Sources - LDH Series", URL `https://www.picoquant.com/products/category/picosecond-pulsed-sources/ldh-series-picosecond-pulsed-diode-laser-heads`, last checked 2021

[100] PicoQuant, "Picosecond Pulsed Drivers - Sepia PDL 828", URL `https://www.picoquant.com/products/category/picosecond-pulsed-driver/pdl-828-sepia-ii-computer-controlled-multichannel-picosecond-diode-laser-driver`, last checked 2021

[101] NKT Photonics, "Super Compact Supercontinuum Lasers", URL `https://www.nktphotonics.com/lasers-fibers/product/superk-compact-supercontinuum-lasers/`, last checked 2021

[102] NKT Photonics, "SuperK Varia Tunable Filter", URL `https://www.nktphotonics.com/lasers-fibers/product/superk-varia-tunable-single-line-filter/`, last checked 2021

[103] J. Lidgard, "The Scattering Calibration System and Anti-Neutrino Detection Studies for SNO+", Ph.D. thesis, University of Oxford (Not yet submitted)

[104] S. Langrock, "Measurement of the Rayleigh scattering length and background contributions during early data taking phases at SNO+", Ph.D. thesis, Queen Mary University of London (2017)

[105] T. Tunstall, Private communication (2017)

[106] J. Dunger, "Multi-hit Corrections for SMELLIE and Energy Reconstruction", SNO+ DocDB 3751-v1 (2016)

[107] A. Zummo, Private communication (2019)

[108] B. Liggins, "CSS update", SNO+ DocDB 5270-v2 (2018)

[109] J. R. Wilson, "An LED Event Generator for RAT", SNO+ DocDB 222-v4 (2010)

[110] I. Coulter, "Photon thinning factor in RAT", SNO+ DocDB 1309-v1 (2012)

[111] E. Hecht, "Optics [electronic resource]", Pearson, Harlow, England, 5th (global) edn. (2017)

[112] M. Wurm et al., "Optical scattering lengths in large liquid-scintillator neutrino detectors", Review of Scientific Instruments vol. 81 (5) (2010)

[113] C. Ransom, "An Analytic Method for Measuring Scattering in the SNO+ Detector", Master's thesis, University of Oxford (2016)