Solar neutrino detection: CNO discovery with Borexino and preparations for success in JUNO and OSIRIS

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Alexandre S. Göttel

aus

London, Großbritannien

Berichter: Univ.-Prof. Dr. Dr. Livia Ludhova
Univ.-Prof. Dr. Achim Stahl

Tag der mündlichen Prüfung: 18.10.2022

Alexandre S. Göttel

Solar neutrino detection: CNO discovery with Borexino and preparations for success in JUNO and OSIRIS

April 2022

Reviewers: Prof. Dr. Dr. Livia Ludhova and Prof. Dr. Achim Stahl

Supervisor: Prof. Dr. Dr. Livia Ludhova

RWTH Aachen University

Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen

III. Physikalisches Institut B

Sommerfeldstr. 14

D-52074 Aachen

You are what you know. Fifteenth-century Europeans 'knew' that the sky was made of closed concentric crystal spheres, rotating around a central earth and carrying the stars and planets. That 'knowledge' structured everything they did and thought, because it told them the truth. Then Galileo's telescope changed the truth.

James Burke, *The Day the Universe Changed*, 1986

Abstract

The Sun is powered through different fusion processes that can be grouped in two categories: the *pp*-chain and the CNO cycle. Building a theory of the balance between gravitational and radiation forces is difficult, not least because almost all of the information about the Sun can only be gathered at a surface level. Particle physics however predicts that the two aforementioned processes emit different kinds of neutrinos: *solar neutrinos*. While their elusive nature makes them extremely difficult to detect, the same property allows them to escape from the solar core. They are the only direct probe of the solar core's internal processes and their measurement on Earth represents a tremendous experimental and theoretical success. This thesis explores methods related to the experimental challenges of measuring CNO neutrinos. These neutrinos constitute less than one percent of the Sun's neutrino output but their generating process is thought to be the main source of stellar energy in our universe. The findings in this thesis led in part to the first experimental evidence for CNO neutrinos with the Borexino detector, which was published in Nature. Another part of this thesis focuses on JUNO's expected sensitivity to solar neutrinos. JUNO is a very large liquid scintillator currently under construction in south China and is shown in this thesis to have the potential for unprecedentedly precise solar neutrino measurements. These sensitivity studies are also currently under preparation for publication. Finally, this thesis encompasses work with the OSIRIS detector, a pre-detector for JUNO which will make sure its liquid scintillator doesn't exceed certain levels of contamination - which could dismantle its sensitivity not only to its main goal of measuring the neutrino mass hierarchy with 3σ in six years, but also severely hinder its solar neutrino measurements. The first part of this section is about developing a source calibration program which will perform all of OSIRIS's calibration needs: energy reconstruction, vertex reconstruction, charge reconstruction, and inter-PMT time offsets on a sub-nano-second scale. It will also be used to monitor OSIRIS's scintillator for time-dependent changes thus providing fast feedback. The results presented in this thesis are also currently being implemented in a paper which is planned to be published once the first calibration data has been measured. In the second part of this section, OSIRIS's sensitivity to its main goal of measuring ^{238}U and ^{232}Th in the liquid scintillator is calculated with an improved analysis compared to existing ones. The results on OSIRIS's sensitivity to its main goal and to ^{85}Kr were already published in the European Physical Journal C as part of an overall design and sensitivity review.

Zusammenfassung

Die Sonne wird durch verschiedene Fusionsprozesse angetrieben, die in zwei große Kategorien eingeteilt werden können: die *pp*-Kette und der CNO-Zyklus. Es ist schwierig, eine Theorie über das Gleichgewicht zwischen Gravitations- und Strahlungskräften aufzustellen, nicht zuletzt deshalb, weil fast alle Informationen über die Sonne nur von dessen Oberfläche gesammelt werden können. Die Teilchenphysik sagt jedoch voraus, dass die beiden oben genannten Prozesse unterschiedliche Arten von Neutrinos aussenden: *Solare neutrinos*. Während sie aufgrund ihrer schwer fassbaren Eigenschaften extrem schwer nachzuweisen sind, ermöglicht ihnen dieselbe Eigenschaft, aus dem Sonnenkern zu entkommen. Sie sind die einzige direkte Sonde aus dem Sonnenkern, und ihre Messung auf der Erde stellt einen enormen experimentellen und theoretischen Erfolg dar. In dieser Arbeit werden Methoden untersucht, die sich mit den experimentellen Herausforderungen bei der Messung von CNO-Neutrinos befassen. Diese Neutrinos machen weniger als ein Prozent der Neutrinostrahlung der Sonne aus, aber man nimmt an, dass ihr Erzeugungsprozess die wichtigste Quelle stellarer Energie in unserem Universum ist. Die in dieser Arbeit gewonnenen Erkenntnisse führten zum Teil zum ersten experimentellen Nachweis von CNO-Neutrinos mit dem Borexino-Detektor, der in Nature veröffentlicht wurde. Ein weiterer Teil dieser Arbeit befasst sich mit der erwarteten Empfindlichkeit von JUNO für solare Neutrinos. JUNO ist ein sehr großer Flüssigszintillator-Detektor, der derzeit in Südchina gebaut wird. In dieser Arbeit wird gezeigt, dass es das Potenzial für beispiellos präzise Messungen solarer Neutrinos hat. Diese Empfindlichkeitsstudien werden derzeit ebenfalls zur Veröffentlichung vorbereitet. Schließlich befasst sich diese Arbeit auch mit dem OSIRIS-Detektor, einem Vordetektor für JUNO, der sicherstellen soll, dass der Flüssigszintillator von JUNO ein bestimmtes Maß an Verunreinigung nicht überschreitet, was nicht nur das Hauptziel der Messung der Neutrinomassenhierarchie mit 3σ in sechs Jahren gefährden, sondern auch die Messungen von Sonnenneutrinos stark beeinträchtigen könnte. Dieser Abschnitt der Dissertation befasst sich erstmalig mit der Entwicklung eines Kalibrierungsprogramms, das alle Anforderungen von OSIRIS erfüllt: Energierekonstruktion, Positionsrekonstruktion, Ladungsrekonstruktion, und Zeitversatz zwischen PMTs im Sub-Nanosekundenbereich. Es wird auch dazu verwendet, den Szintillator von OSIRIS auf zeitabhängige Veränderungen zu überwachen und so ein schnelles Feedback zu liefern. Die in dieser Arbeit vorgestellten Ergebnisse werden derzeit auch in einer Publikation umgesetzt, die veröffentlicht werden soll, sobald die ersten Kalibrierungsdaten gemessen wurden. Weiterhin wird in dieser Arbeit die Empfindlichkeit von OSIRIS für dessen Hauptziel, die Messung von ^{238}U und ^{232}Th im Flüssigszintillator berechnet. Die Ergebnisse dieser Studie, sowie Ergebnisse zur Empflichdichkeit zu ^{85}Kr im Flüssigszintillator, wurden bereits im European Physical Journal C als Teil einer allgemeinen Review von OSIRIS veröffentlicht.

Contents

Introduction

This thesis is organised in six sections. The first two focus on neutrinos and the detectors that were used and/or studied in my work, and the last four are about my main contributions to the Borexino and JUNO collaborations. Section 1 gives an introduction to neutrinos as a particle, to their 100-year old history, to the physical laws governing their interactions, and how they can be detected through interactions with liquid scintillator. It describes different kinds of neutrino sources, with a particular emphasis on neutrinos from the Sun, hereafter referred to as *solars*. It also introduces the concept of CNO neutrinos, which are neutrinos from a rare fusion sub-process in the Sun called the CNO cycle. In stars whose mass is at least 1.3 times that of the Sun, the CNO cycle is thought to be the driving process behind helium fusion, therefore making it the most important source of stellar energy in our universe.

Section 2 introduces Borexino: a 300 t liquid scintillator detector that has taken data from early 2011 to late 2021 and has led to measurements of all pp-chain neutrinos (except the so-called *hep* neutrinos) and to the first experimental evidence for the existence of CNO neutrinos - which is one of the focal points of this thesis. It also introduces the JUNO experiment which is a 20 kt liquid scintillator currently under construction in the Guangdong province is south China. JUNO's enormous target mass will allow it to accumulate immense numbers of solar neutrinos and will therefore be able to measure their rate with unprecedented accuracy. This accuracy is however greatly dependent on the internal radioactivity of the liquid scintillator, whose so-called *radio-purity* is also of great importance (to JUNO and to this thesis). Finally, it introduces the OSIRIS detector which is a pre-detector for JUNO and another major focus of this thesis. OSIRIS will monitor the liquid scintillator after its purification at the JUNO site but before its entry to JUNO's central volume. By being able to give quick alarms or even trigger vetoes, OSIRIS will ensure that the contamination of the liquid scintillator through uranium and thorium will not exceed about $1 \times 10^{-15}\,\mathrm{g\,g^{-1}}$.

Section 3 introduces the analysis methods which were used in Borexino in order to perform measurements of solar neutrinos. It then describes the strategies that were used to detect CNO neutrinos in Borexino, with an emphasis on the most nefarious background: ^{210}Bi in the liquid scintillator. In section 3.6 my contribution to the analysis is delved into: providing an independent constraint on ^{210}Bi. It explains how I designed a modern Bayesian framework based around nested sampling in order to set an upper limit on the background component. The framework, which I implemented with Sindhujha Kumaran, uses spline-based functions in order to model the spacial distribution of the so-called "low-polonium-field", a region of minimal background noise, while making use of a Bayesian factor analysis to make sure the results are as unbiased as possible. It was

also employed to monitor the time-dependence of the region which was found to move up and down. My method proved to be much more stable than previous more traditional ones, could straightforwardly implement physical constraints through Bayesian priors, and gave us the possibility to quantitatively test different models on the available data. Finally, section 3.7 explains how this constraint was used in order to provide the very first experimental evidence for the existence of CNO neutrinos (and of this process occurring in our Sun), which was published on the front page of Nature.

Section 4 serves to quantify the expected precision with which JUNO will be able to measure the rates of ^{7}Be, *pep*, and CNO neutrinos. This is highly dependent on the radioactivity levels that will be achieved in JUNO, but also on the performance of background reduction techniques. It describes all different backgrounds that can affect the results: namely *internals*, *externals*, and *cosmogenics*. It also describes a piece of software, developed in C++ using ROOT::TMinuit, which I designed and built with the help of my colleagues. The software, called JUST (Jülich nUsol Sensitivity Tool), is able to perform quasi-arbitrary quantities of toy-data fits on spectra which are Poisson-fluctuated on the fly. The results of these fits on different parameters directly give the sensitivity of JUNO to solar neutrinos in different scenarios. My work was not only being the main developer of JUST, but also estimating the levels of the different backgrounds that are expected JUNO, and finally using JUST to study JUNO's sensitivity as a function of different background contributions.

Section 5 details the largest portion of my research: calibration studies for the OSIRIS detector. As mentioned above, OSIRIS will monitor the radio-purity of JUNO's liquid scintillator before it flows into JUNO. It therefore plays an important role not only in JUNO's main analyses concerning the mass hierarchy, but also in setting limits on (and measuring) internal backgrounds for JUNO. This is particularly relevant for the solar analysis discussed in section 4 as its results are highly dependent on these backgrounds. OSIRIS's calibration will be performed with two systems: *laser* and *source*. My work was focused on the design and optimisation of the source system. It will be used to lower three different sources directly inside of OSIRIS's liquid scintillator. The first source contains three radioactive gamma isotopes (^{137}Cs, ^{65}Zn, and ^{60}Co) and will be used to calibrate OSIRIS's energy response. The second, containing low-activity ^{40}K, will be submerged inside of liquid scintillator during normal physics runs and provide a reference peak to monitor time-dependent changes in the scintillator. These have been known to occur, for example in Daya Bay, and in the case of OSIRIS which will constantly be measuring new scintillator flowing through it, will not only provide quick alarms in case of problems but also provide invaluable feedback to the team performing the purification of the scintillator (the last step in the scintillator production which occurs just before it flows into OSIRIS). Lastly, one source will contain a fast-pulsed LED capable of emitting

a low number of photons per pulse. This will serve, in redundancy to the laser system, to calibrate the time differences between PMTs on a nano-second scale and to perform a s.p.e. charge calibration of the PMTs. Because the intensity of the LED pulses can be changed on-the-fly, it will also be used to study the charge-dependence of the time responses of the PMTs.

Finally, section 6 presents the work I performed in the estimation of OSIRIS's sensitivity to different JUNO-relevant backgrounds. A full simulation and complex calculation was performed in order to estimate efficiencies related to so-called Bi-Po decays, which are signatures of ^{238}U and ^{232}Th in the detector. The statistical treatment is then outlined and the results presented. The results of a more realistic calculation are then also discussed to display the effect of $^{222-220}$Rn contaminations in the detector on the sensitivity. All-in-all it is shown that OSIRIS will be able to reach the aforementioned levels of 1×10^{-15} gram/g, but only if the radon contamination can be measured with at least 10% precision. Similar calculations are then presented in the context of the search for a rare decay of ^{85}Kr. It is shown that although OSIRIS will not be sensitive to expected background ^{85}Kr levels, it will be able to detect it in case of and air leak in the scintillator. ^{85}Kr signatures will in that case be very useful as they are a definitive proof that an air leak has occurred. Finally, OSIRIS's precision to measuring ^{14}C and ^{210}Po is presented.

1 Neutrino physics

This section gives an overview of what a neutrino is, of the fundamental properties that govern it, and of the current status of neutrino research. As such, section 1.1 gives a historical introduction to neutrino physics, section 1.2 explains the principles of neutrino oscillations, which are necessarily at the core of neutrino physics. Then, sections 1.3–1.5 elaborate on different neutrino sources (with a focus on *solar* neutrinos which are the most relevant for this thesis, and on *reactor* neutrinos which are the main object of the JUNO detector on which this thesis partly focuses), section 1.6 describes how neutrinos are detected using liquid scintillator detectors, and finally, section 1.7 lists the open questions currently most prevalent in the field of neutrino physics.

1.1 History

The history of neutrino physics was naturally born out of physical considerations which, as we now know, led to a rich new field and the genesis of many new open questions. Since many of the questions of the time are still (or again) active research areas today, this section gives a concise introduction to neutrino physics and the mysteries that still accompany it by following a historical series of events. The first such event is more of an unexpected experimental debacle which in the 1920s led to many (probably heated) discussions about the observed energy spectrum of beta decay electrons. Instead of the sharp peak that was expected at a theoretical Q-value, as expected from a two-body decay, a continuous spectrum with a maximum energy of exactly the theorised Q-value was measured in several experiments. This seemed to break energy conservation, spin conservation, and even the observed direction of the momentum of the particles couldn't be explained. In an attempt to settle the arguments, Ellis & Wooster designed an experiment that was performed in 1927 [1]. The experiment used a Radium source enclosed in a thick calorimeter (large enough to block any *known* particle with those energies). They observed through the heating of the calorimeter that the average energy of the events was compatible with the previously measured average energy of the beta-decay electrons. This proved that the electrons were emitted from the atoms with a continuous energy distribution and that no "secondary effect" (outside of the nucleus) was the cause of the continuous distribution.

With a theoretical explanation of the continuous spectrum still lacking, this ultimately led to Pauli famously proposing the existence of the neutrino in 1933. This was quickly picked up by Fermi who named the particle and developed a quantitative theory of nuclear decay based around it [2]. At this point, an upper limit on the mass of the neutrino

was already set at around 1 keV. This upper limit was based on tritium-endpoint measurements, much like the Katrin experiment [3] is doing today! Great efforts were put in the (originally thought to be impossible) detection of neutrinos in the next years. In 1952, Rode & Allen [4] achieved an indirect confirmation of the neutrino hypothesis by analysing the recoil of ^{37}Cl ions from ^{37}Ar decays - whose energy was found to be entirely consistent with the hypothesis. After a (perhaps daring, perhaps desperate) proposal for an experiment meant to measure neutrinos from nuclear weapon explosions which was thankfully cancelled, Reines & Cowan built a large liquid scintillator detector (three volumes of 200 l each) which could detect antineutrino-induced inverse beta decays via the positron-neutron coincidence (see section 1.6) in 1956 [5]. Using reactor neutrinos as a source and Cadmium as a neutron target, they found a highly significant neutrino event rate of $(2.88 \pm 0.22)\,\mathrm{h}^{-1}$, thereby directly proving the existence of the particle. The next decades saw an explosion in the research performed in neutrino physics as well as in the variety of important and significant results in the field. In 1962, neutrinos originating from Brookhaven's Alternating Gradient Synchrotron were used as the first neutrino beam to prove the existence of muon neutrinos [6]. This was followed by another neutrino beam being used in 2001 by the DONUT collaboration that proved the existence of tau neutrinos [7]. While there are now more neutrino experiments than is sensible to list here, the ones that are most relevant for this thesis are discussed below.

The most fundamental question about neutrinos is of course their particle properties. Already from the time of the proposal of neutrinos by Pauli, it was known that neutrinos must be neutral in all but the weak force - since they hadn't been detected yet and because of their obvious origin from nuclear decay. It was further determined that neutrinos are fermions with a spin of 1/2 and a weak isospin of -1/2. It was also very quickly clear from the aforementioned beta electron spectrum that their masses, should they have any, must be very little. In the decades following the discovery of the neutrino the existence of more than one family of neutrinos was observed. The question of course quickly came up of how many of those neutrino families can exist. A measurement of the Z-decay resonance at LEP [8] showed that the number of light "invisible" particles that can interact with the Z-boson (and whose mass is less than half the mass of the Z-boson) is exactly three, which corresponds to the known number of neutrino families. In theory, additional so-called *sterile* neutrinos could exist which don't even interact weakly, and there are several hints for their existence (as well as good arguments against it), see section 1.7.

It was further observed by Goldhaber, in 1958, that the helicity of neutrinos is always negative [9]. This was done by observing the polarisation of gammas from the electron capture decay of ^{152m}Eu. Consequently, in the Standard Model of particle physics, neutrinos must be massless as in any other case one would expect some probability

for right-handed neutrinos to be observed. The aforementioned observation of neutrino oscillations, which necessarily implies non-zero neutrino masses, constitutes (to this day the only) strong evidence of physics beyond the Standard Model. Most of the other possible properties are currently unknown or in the process of being discovered. Section 1.7 gives an overview of the most important ones.

Of course at the same time the theory of the neutrino significantly expanded as well. Many ideas were floating in the scientific community, including the idea of neutrino mixing (similar to quark mixing). This led to the idea of neutrino oscillations, by Pontecorvo in 1958 [10, 11], which implied that neutrinos undergo flavour conversion when travelling. This conversion was proven to exist by the SuperKamiokande experiment [12] and the SNO collaboration [13] in 1998 and 2001, respectively. Their observations also solved the so-called *solar neutrino problem* [14] which referred to a significant deficit in the measured rates of solar neutrinos compared to theoretical observations (that did not account for flavour conversion).

This deficit was first observed in the Homestake experiment with only 30% of the expected rate measured (using a Chlorine-based radiochemical detection method that exploits CC ν_e reactions, see section 1.6.5) [14]. This was followed up by SAGE who in 1994 reported a 42% deficit [15] (the difference in deficits is due to a difference in energy thresholds), which was confirmed by the Gallex and GNO experiments [16, 17] which both used the same Gallium-based radiochemical methods. In 1998, SuperKamiokande also measured a deficit (of 59%) using a completely different detection method [12]. The problem was resolved by SNO [18] which exploited the interactions of ^{8}B neutrinos (see section 1.3.2) on deuterium oxide and was able to separate neutrinos that interacted through different kinds of interactions. It was indeed able to differentiate between signals from neutrinos that had undergone neutral-current (NC), charged-current (CC), or elastic-scattering (ES) interactions (see section 1.6.5). The flux from NC interactions, in which all neutrino flavours can participate, was exactly that expected from the SSM. The fluxes from CC interactions, in which only ν_e solar neutrinos can participate, and the fluxes from ES interactions where all flavours are involved but where ν_e have a six times larger cross-section both showed a deficit of around 70%. The results can be seen in fig. 1. This proved that flavour conversion of solar neutrinos does in fact occur. The final proof that oscillation is the process behind this was finally given by KamLAND in 2008 [19] which performed the first observation of a full oscillation period. The next sections go into more detail of neutrino oscillations, which turned out to basically be an entire field of research on their own.

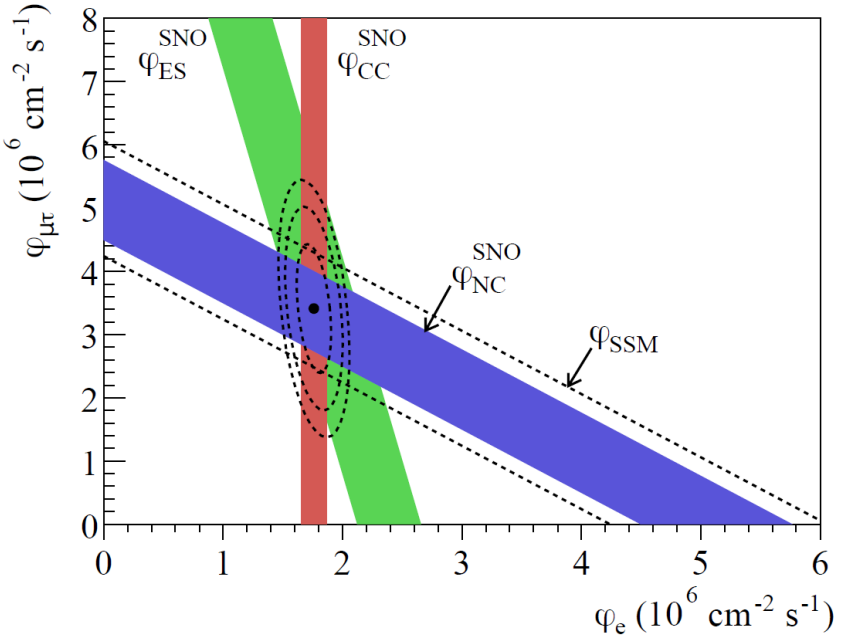


FIGURE 1: Results from the SNO experiment [13]. The solid bands show the flux measurements from the CC, ES, and NC reactions. The dotted band shows the SSM expectation. The bands clearly intersect in the interval from a fit results - showing that there is a good agreement between model and measurement.

1.2 Neutrino oscillations

As previously mentioned, the theory of neutrino oscillations was first put forward by Pontecorvo in 1958 [10]. The idea is based on the possibility of non-zero neutrino masses, which could induce a mixing phenomenon that is parallel in many ways to mixing in the quark sector [20]. At its core, it is an assumption of the existence of a non-trivial unitary transformation between the flavour eigenstates of neutrinos, *i.e.* the states that interact with weak bosons, and its mass eigenstates. This takes the following form:

$$\begin{aligned} |\nu_\alpha\rangle &= \sum_j \mathcal{U}^\star_{\alpha j} \, |\nu_j\rangle \\ |\nu_j\rangle &= \sum_\alpha \mathcal{U}_{\alpha j} \, |\nu_\alpha\rangle \, , \end{aligned} \tag{1}$$

where $\alpha \subset \{e, \mu, \tau\}$ denotes the flavour eigenstate associated with the respective leptons, $j \subset \{1, 2, 3\}$ denotes the mass eigenstate, and $\mathcal{U}$ is the so-called *Pontecorvo-Maki-Nakagawa-Sakata* (PMNS) matrix [11, 21]. When propagating a mass eigenstate using a plane wave approximation and assuming an ultrarelativistic regime (always well-funded for neutrinos), one quickly obtains:

$$|\nu_k(L)\rangle = e^{-i\frac{m_k^2 \cdot L}{2E}} \, |\nu_k(0)\rangle \, , \tag{2}$$

where m_k is the mass of the corresponding state, L is the distance travelled, and E is the energy. One can then straightforwardly calculate the probability of the state of a neutrino which was emitted in the state α to be measured in the state β:

$$\begin{aligned} P_{\alpha\to\beta} &= |\langle\nu_\beta(L)|\nu_\alpha(0)\rangle|^2 \\ &= \sum_{k,j} \mathcal{U}_{\alpha k}\mathcal{U}^\star_{\beta k}\mathcal{U}^\star_{\alpha j}\mathcal{U}_{\beta j} e^{-\frac{i}{2}\cdot\frac{\Delta m^2_{kj}\cdot L}{2E}}, \end{aligned} \tag{3}$$

where $\Delta m^2_{ij} = m^2_j - m^2_i$. One should note that a more detailed quantum mechanical treatment of neutrino oscillations must consider that since the wave packets associated with the different mass eigenstates move with different speeds, an oscillation can only be expected up until the point where those wave packets don't overlap anymore [22]. Because of the extremely low mass squared differences innate to neutrinos, this decoherence was never observed. However, because of the somewhat larger Δm^2_{32} to which reactor neutrino experiments are sensitive (see section 1.4) and because of the small size of atomic nuclei, which constitute the source for reactor neutrinos, some reactor neutrino experiments have been able to set tight constraints on the effect [23, 24]. In the future this can be expected to be improved on by other reactor neutrino experiments, not least with JUNO [25], and such an observation would constitute a powerful macro-scale confirmation of a quantum mechanical effect.

At this point the PMNS matrix can be more thoroughly introduced. In its most common form it takes the shape of three rotations in the flavour-mass space, with an additional phase for possible charge-parity violation:

$$\mathcal{U} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4}$$

where $c_{ij} = \cos(\theta_{ij})$, $s_{ij} = \sin(\theta_{ij})$, δ is a charge-parity violating phase, and $\alpha_{1/2}$ are so-called Majorana phase factors (exist only if neutrinos are Majorana particles [20] but it has no effect on the oscillation patterns, see section 1.7). As is obvious from its form, the PMNS matrix was parametrised with the expectation of a unitary transformation. While the current measurements of the unitarity of the PMNS matrix are limited in precision [26] because of the difficulty of the full measurement (particularly involving ν_τ), it is currently under investigation. Many of the parameters used in neutrino physics rely on assumptions of unitarity.

In order to see the effect of neutrino oscillations more easily than in eq. (3), one can look at what one would obtain for the probability of a flavour change in a two-flavour

limit, inserting the elements from the PMNS matrix:

$$P_{\alpha\to\beta\neq\alpha} = \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2 \cdot L}{4E}\right), \tag{5}$$

where θ is the associated oscillation angle. One can directly here see the oscillating form, as well as the three most important parameters: L/E, the ratio between the distance travelled and the energy (which can be controlled depending on the neutrino source and experiment), the mass square difference Δm for the period of the oscillation, and the angle θ, which is proportional to the amplitude of the oscillation. Indeed, as one can see, should the masses of the two states be equal, then no oscillation could occur. In the three-neutrino paradigm, there are six independent parameters: the angles θ_{12}, θ_{23}, and θ_{13}, the two mass square differences Δm^2_{21} and Δm^2_{32}, and the CP-violating phase δ_{CP} (see section 1.7 for more discussion). The current status of the measurement of these parameters is summarised in table 1.

	Global fit (NH)	Global fit (IH)	Dominant Source
$\sin^2\theta_{13}$	$0.02220^{+0.00068}_{-0.00062}$	$0.02238^{+0.00064}_{-0.00062}$	Reactor MBL
$\sin^2\theta_{12}$	$0.304^{+0.013}_{-0.012}$	$0.304^{+0.012}_{-0.012}$	Solar
$\sin^2\theta_{23}$	$0.573^{+0.018}_{-0.023}$	$0.578^{+0.017}_{-0.021}$	Beam ν_μ disappearance
Δm^2_{21}	$7.42^{+0.21}_{-0.20}10^{-5}\text{eV}^2$	$7.42^{+0.21}_{-0.20}10^{-5}\text{eV}^2$	Reactor LBL
Δm^2_{32}	$2.515^{+0.028}_{-0.028}10^{-3}\text{eV}^2$	$-2.498^{+0.028}_{-0.029}10^{-3}\text{eV}^2$	(very)LBL Beam
δ_{CP}	194^{+52}_{-25}	287^{+27}_{-32}	LBL Beam ν_e appearance

TABLE 1: Global fit results [27, 28] on neutrino data from many experiments (September 2020) when assuming normal and inverted orderings (NH, IH). The third column shows the dominant contribution to the sensitivity of each parameter to the global fit. LBL and MBL stand for "long baseline" and "medium baseline", respectively. It is to be noted that when including data from atmospheric neutrino measurements from SuperKamiokande, the value of $\sin^2\theta_{23}$ goes from 0.573 to 0.450 for the NH, but remains largely unchanged for the IH. Additionally, it is unclear whether this value of $\sin^2\theta_{23}$ is correct since a mirrored value on the other side of 45° is also likely, see section 1.7.

θ_{12} and Δm^2_{21} have been measured by solar neutrino experiments. Particularly sensitive are precise measurements of ^{8}B solar neutrinos, the most powerful ones stemming from SuperK [29] and SNO [30]. Medium/long baseline reactor neutrino experiments are also sensitive to both parameters, and while solar experiments offer the highest precision for θ_{12}, the highest contribution to the global fit for Δm^2_{21} is KamLAND [31]. θ_{23} and Δm^2_{32} can both be inferred quite accurately by atmospheric neutrino measurements, which has been done by Antares, IceCube, and most accurately by SuperKamiokande [32–34]. The most precise measurements for both parameters are from very long baseline accelerator neutrino experiments, most notably T2K [35] but also NOνA [36] and MINOS [37]. Because of its small value, θ_{13} was first determined to be non-zero before being measured

precisely ($\approx 3\%$) by Daya Bay [38], and then also measured by Double Chooz [39] and RENO [40]. Beam experiments are currently mainly limited in their sensitivity to θ_{13} by the θ_{23}-ocant degeneracy, see section 1.7. As oscillation parameters go into an age of precision measurements, JUNO will contribute enormously with expected sub-percent (unprecedented) measurements of Δm^2_{31}, Δm^2_{21}, and $\sin^2\theta_{12}$. Finally, the CP-violating phase δ_{CP} can be measured through the appearance of ν_e in ν_μ beams, which can currently only be done by T2K [35] and NOνA [41]. T2K recently excluded the $\delta_{CP} = 0$ hypothesis with a 95% C.L..

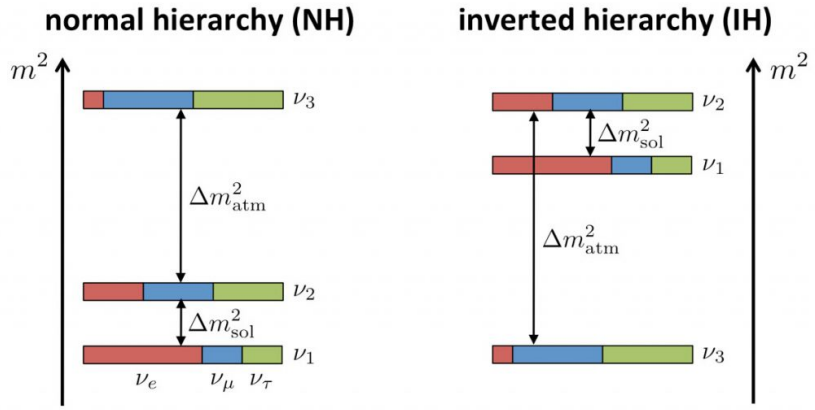


FIGURE 2: Graphical representation of the mass hierarchies. Δm^2_{atm} refers to Δm^2_{32}, and Δm^2_{sol} to Δm^2_{21}. The colours of the different mass states show the probabilities of interacting with red:ν_e, blue:ν_μ, green:ν_τ. The figure is not to-scale. *Credit*: JUNO Collaboration / JGU-Mainz

As one can see from the table, there are two paradigms under which the parameters are evaluated: the normal hierarchy (NH) and the inverted hierarchy (IH). These refer to the fact that the sign of Δm^2_{32} is yet unknown. As such, there are two possibilities for the orderings of the three different mass terms: $m_1 < m_2 < m_3$(normal) or $m_3 < m_1 < m_2$(inverted). The two possibilities are depicted in fig. 2, where one can intuitively understand the existence of both orderings. As a reminder: Δm^2_{21} is about two orders of magnitude smaller than Δm^2_{32}. The *normal* ordering was named as-is because this ordering goes hand-in-hand with the ordering of the corresponding lepton masses. The nature of the ordering is a very important parameter for cosmology and for the planning of future experiments and is currently the subject of a lot of research, see section 1.7.

Because of the prevalence of electrons in normal matter and the innate interactions between electron neutrinos and electrons, the probabilities for oscillation can change significantly when neutrinos are travelling through matter. This was first theorised by Wolfenstein in 1978 [42] then expanded on by Mikehyev and Smirnov in 1986 [43]. By analysing the effect of mostly coherent scattering of neutrinos on electrons in matter,

they found that the following correction must be applied on oscillations (here for simplicity assuming again a 2-flavour framework) which can be done by introducing a modified oscillation angle:

$$\sin^2 2\theta_m = \frac{\sin^2 2\theta}{(2\sqrt{2}G_F N_e E/\Delta m^2 - \cos 2\theta)^2 + \sin^2 2\theta}, \tag{6}$$

where θ_m is the *modified* angle θ, G_F is the Fermi constant, N_e is the electron density in the matter, and E is the energy of the neutrino. One can directly see from the formula that the modified angle goes into a resonance when:

$$E = \frac{\Delta m^2 cos 2\theta}{2\sqrt{2}G_F N_e}. \tag{7}$$

A full calculation of the effect of matter oscillations in the Sun [44] shows that at energies much lower than the resonance values the survival probability of electron neutrinos converge towards the *vacuum* solution, and that for higher energies it eventually converges towards a lower value, see fig. 3. The most precise measurement of this effect in the Sun was done by SuperKamiokande [29] by measuring the flux of ^8B solar neutrinos. While experimental evidence for the MSW effect has been conclusive since 2004 [45], data from solar neutrinos serve as an important confirmation. The next big step in this area was from the Borexino experiment with the only observation of all pp-chain solar neutrinos [46] therefore covering neutrinos above and below the transition region in full compatibility with the predictions from the MSW effect, see fig. 3.

In practice one could also observe the MSW in the Earth by looking at the day-night asymmetry of ^{8}B solar neutrinos. SuperKamiokande has reported an asymmetry of $-3.3 \pm 1.0(stat) \pm 0.5(syst)$ [29] which rejects the non-MSW hypothesis with 2.9σ, but a recent update on their data indicates that the significance is in fact even smaller [47].

1.3 Solar neutrinos

1.3.1 The standard solar model (SSM)

The goal of this section is to give a brief overview of the SSM and of the recent progress in the theory, while not going into details as most are beyond the scope of this thesis.

The SSM [48] is the best currently available model of our Sun. It was firstly developed in the 80s, where it was set up as a tool to satisfy basic structural equations known to be valid for any star. These equations include hydrostatic equilibrium, the balance between inward gravitational and outward radiative forces, and the transport of energy by photons and by convection, separated in different layers.

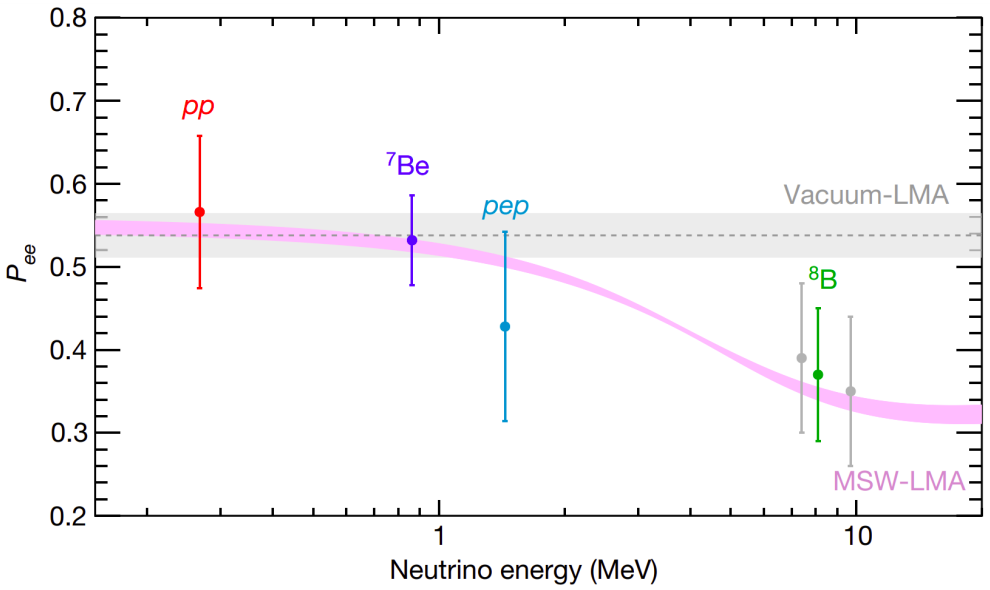


FIGURE 3: Solar electron neutrino survival probability as a function of energy. The grey band shows the 1σ uncertainty band when no matter effects exist, and the pink one when they do. The data points are based on different kinds of solar neutrino measurements from Borexino, and the grey points show different sub-energy-regions for ^{8}B neutrinos. As one can see, Borexino's data is compatible with the MSW solution. *Credit*: Borexino collaboration [46]

The SSM is based on several physical input quantities, such as the initial mass and the ratio of Helium to heavier elements at the time of the conception of the star. These inputs also include the current radius and luminosity of the Sun. As an output, the SSM produces measurable quantities, including the abundances of different elements in the surface layer, sound speed profiles, and the rate of neutrino production in the core.

In the past decades, the SSM has been a stable ground used to add more-and-more refined models of the different processes, as well as more detailed and precise inputs. The more basic SSMs were already enough to provide excellent agreement with SNO neutrino observations, which led to the 2002 Nobel prize to R. Davis and M. Koshiba [13] after solving of so-called "solar neutrino problem" (see section 1.1). There are however still a lot of unknowns to understand and more details to add. For example, the plasma equation of state was recently extended to include relativistic electronic corrections [48]. To give an idea of the complexity of the solar interior, a sketch of the current understanding can be seen in fig. 4.

More importantly, in the context of this thesis, recent updates in the observed carbon, nitrogen, and oxygen abundances in the outer layers of the Sun, as well as a refined three-dimensional modelling (as opposed to two-dimensional) led to some disagreement with other observables. This is referred to as the *solar metallicity problem*. Indeed, while this update is well-grounded [50] and does lead to the solution of an anomaly

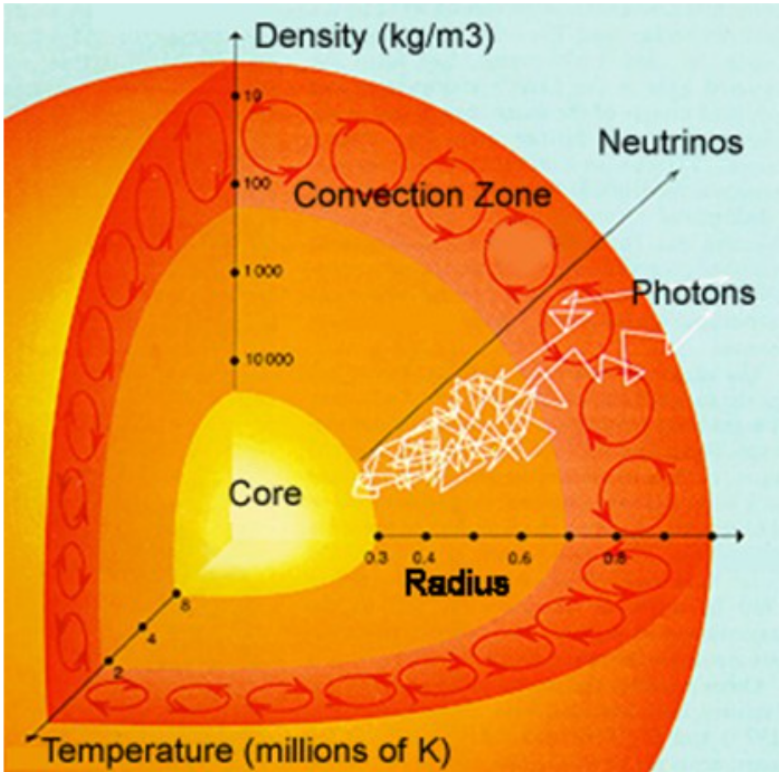


FIGURE 4: Theoretical schematic of the inner workings of the sun [49].

related to oxygen abundances [51], it does also lead to a deteriorated agreement with sound speed profile observations compared to an older model with higher metallicity [52]. Solar neutrino fluxes do in general indirectly depend on the metallicity. This is because the fluxes depend on the core temperature of the Sun T_C, and because the temperature depends on the radiative opacity of the Sun. The uncertainties related to this radiative opacity cause a very high degeneracy, for example because deviations from nominal values can mimic neutrino fluxes that would be expected from a higher metallicity [53–55].

CNO neutrinos are neutrinos from the theorised Bethe-Weizsäcker cycle [56, 57] in the Sun and will be discussed in more detail in section 1.3.3. In contrast to the other solar neutrino fluxes (particularly ^{7}Be and ^{8}B), because CNO reactions depend on the abundances of carbon, nitrogen, and oxygen in the Sun, the flux of CNO neutrino directly depends on metallicity. Following the approaches in [58–60], one can write these dependencies as $\Phi_i \propto T_c^{y_i}$, where $i \subset \{CNO, ^{8}B\}$ and y_i is a constant. Therefore, $\frac{\Phi_{CNO}}{\Phi^{k}_{^{8}B}}$ is practically independent of T_C and thus of the radiative opacity when $k = y_{cno}/y_{^{8}B}$. When used in combination with precise ^{8}B neutrino measurements, for example from SuperKamiokande [29], a measurement of the rate of CNO neutrinos can, following the approaches in [58], be used to infer the abundance of carbon and nitrogen in the Sun. One can thus equate the measurement of ^{8}B neutrinos as a “thermometer” to constrain

T_C. The approach involves expressing $\frac{\Phi_{CNO}}{\Phi_{^8B}}$ as a function of all SSM inputs. When reversing the equation, one obtains [60]:

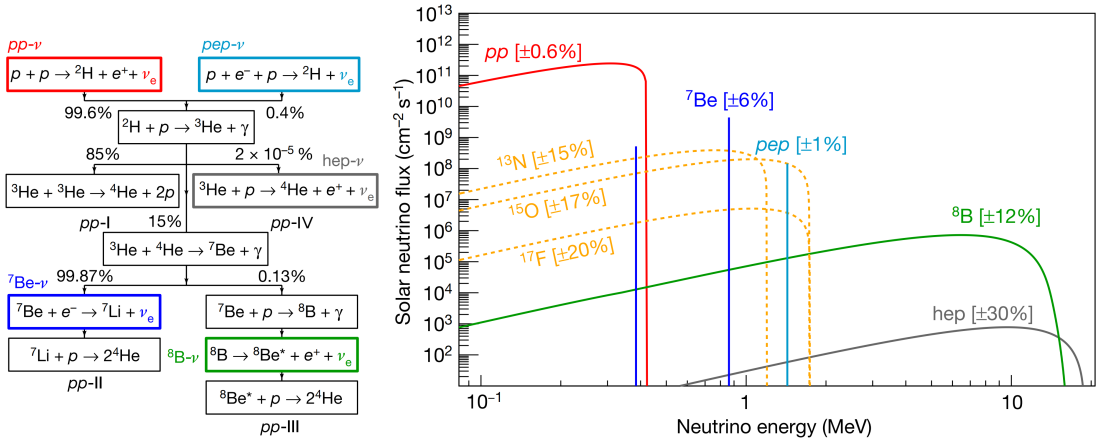
$$\frac{N_C + N_N}{N_C^{SSM} + N_N^{SSM}} = (\frac{\Phi_{^8\mathrm{B}}}{\Phi_{\mathrm{CNO}}})^{-k} \cdot \frac{R_{CNO}^{exp}}{R_{CNO}^{SSM}} \cdot [1 \pm 0.5\%^{(env)} \pm 9.1\%^{(nucl)} \pm 2.8\%^{(diff)}], \quad (8)$$

where N_i is the number density of i with respect to hydrogen, *SSM* denotes SSM predictions, Φ_i is the flux of i, R_i is the interaction rate of i in a given detector, R_i^{exp} is the measure interaction rate of i. Φ_{CNO} only refers to the ^{15}O and ^{13}N contributions, which is a good approximation because that of ^{17}F is very small in comparison. The uncertainties were classified in three different SSM input categories related to the *environment* (the parameters which affect the Sun's temperature profile the most), *nuclear* cross-section, and *diffusion* parameters. Additionally, the measurement of the rate of CNO neutrinos is crucial to understanding the interactions in the radiative zone of the Sun where different models of micro- and macroscopic interactions lead to different CNO compositions in the core.

1.3.2 *pp*-chain neutrinos

There have been suspicions that proton fusion was the source of the Sun's energy production since the 1920s. It was however not before the suggestion of a chain of nuclear reactions by Gamov & Weizsäcker between 1929 and 1937 [57, 61], made possible by developments in quantum mechanics, that the idea took a more concrete formalism. In 1939, Bethe then calculated the different rates of the reactions in a more complete model of the Sun [56, 62]. These reactions have then been studied in more detail since the 50s [63], which led to the idea of using solar neutrinos to study the Sun, and today our understanding of them is quite detailed, see fig. 5.

The first detection of solar neutrinos was performed by the pioneering radiochemical Homestake experiment which started in 1970 [64]. This was followed by similar experiments such as Gallex/GNO [16, 17] and Sage [15] in the 90s, which could all only be sensitive to an integral flux of solar neutrinos. In parallel, other detection methods were successfully developed, such as Kamiokande-II which, using a water-cherenkov detector, led in the 80s to the first real-time measurement of ^{8}B neutrinos [65]. This was quickly followed by SNO in 2001 [13], see section 1.1. In 2007, the Borexino experiment used a liquid scintillator technique in order to measure ^{7}Be solar neutrinos in real time. The measurement of other neutrinos was later made possible through detector and analysis improvements, which culminated in 2018 in the first full simultaneous spectroscopic measurement of all pp-chain solar neutrino (except *hep*).



(a) (b)

FIGURE 5: (a) Summary of all the pp-chain reactions and of its four branches. The coloured boxes show the origin of the respective neutrinos: *pp* in red, *pep* in cyan, ^{7}Be in blue, ^{8}B in green, and *hep* in grey. One can clearly see that the pp-I and pp-II branches are the most important, and the pp-IV chain the least. (b) Spectra & rates of the different solar neutrinos as a function of energy, along with the theoretical uncertainties (68% C.L.) in square brackets [46]. CNO neutrinos are depicted as yellow dotted lines. The ^{7}Be and *pep* fluxes are shown as lines because their two-body decays infer a monochromatic energy distribution.

1.3.3 CNO neutrinos

The CNO cycle is one of the fusion processes through which the conversion of hydrogen to helium in main sequence stars (on the Hertzsprung–Russell diagram[1]) occurs, alongside the *pp*-chain (see section 1.3.2). It is also known as the Bethe-Weizsäcker cycle because the theoretical process was simultaneously discovered by both Hans Bethe and Carl F. Weizsäcker in the late nineteen-thirties [56, 57]. At the time, this cycle was thought to be the primary source of energy production in the Sun. Through later advances in stellar physics, it was understood that although the dependence on temperature of the CNO reactions is much stronger than that of the *pp*-chain, the CNO cycle can only exist in temperatures exceeding 15×10^{6} K [66] (very close to the actual temperature of the solar core[2]), compared to 4×10^{6} K for the *pp*-chain. Therefore, while it is theorised to only contribute a small part of the energy (and neutrino) production in the Sun (1% to 2%), it quickly becomes the dominant process in stars heavier than about 1.3 $M_{\odot}$ ($M_{\odot}$ is the mass of the Sun) [66] or in late evolutionary stages [67]. Figure 6 gives an overview of the main CNO cycle, which effectively converts four protons into a ^{4}He nuclei. Considering that the two positrons emitted by the reaction quickly annihilate in the electron-dense solar medium, the CNO cycle emits on average about 25 MeV of energy in photons and 1.7 MeV in neutrinos.

[1] https://www.e-education.psu.edu/astro801/content/l4_p6.html
[2] https://solarscience.msfc.nasa.gov/interior.shtml

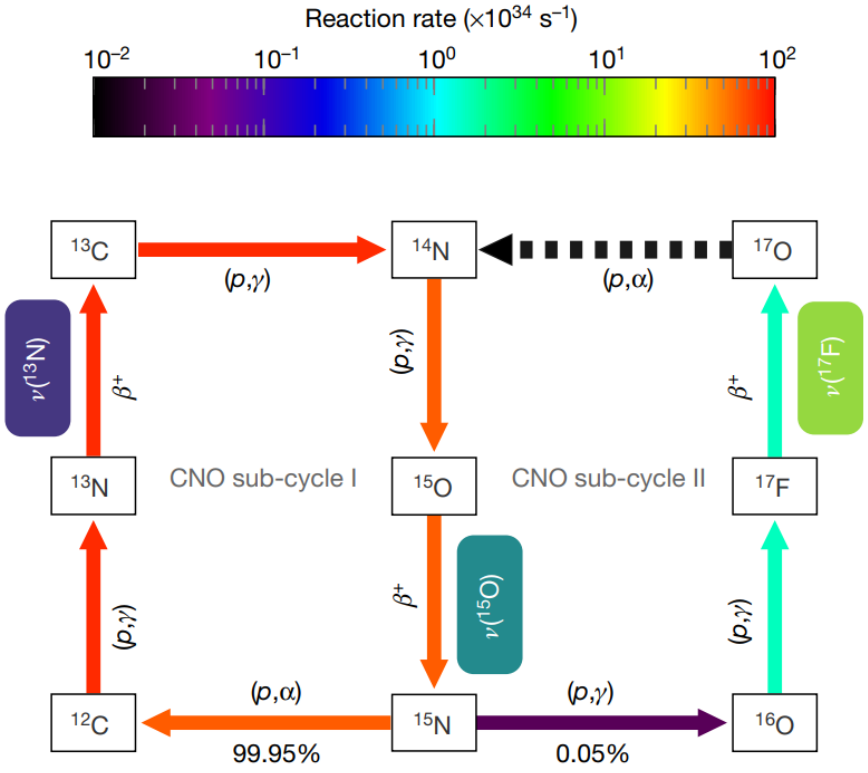


FIGURE 6: The CNO cycle [68]. Starting with ^{12}C, a series of proton captures and beta decays transform the original nucleus until the resulting ^{15}N nucleus α-decays back into the original ^{12}C. One execution of this primary chain effectively fuses four protons into one helium nucleus, thereby releasing about 26.7 MeV of energy. This includes annihilation photons resulting from the positron emissions. The right side of this drawing shows the secondary CNO chain. The colours indicate the reaction rates.

Being the main source of energy production for massive stars ($M_{star} \geq 1.3\ M_{\odot}$), there are different signatures for the existence of the CNO cycle in the universe. For example, while one iteration of the CNO-chain effectively doesn't alter the abundances of C, N, and O, as can be seen in Figure 6, the relative abundances of those elements are affected by the different lifetimes of the reactions. In this case, the $^{14}N(p,\gamma)^{15}O$ reaction inside of the CNO chain has a lifetime that is longer relative to the other reactions, thereby creating a bottleneck which ends up decreasing the $^{12}C/^{13}C$ ratio and increasing abundances of ^{14}N [69]. Thus, analyses of the chemical composition of a star can include clear signatures of CNO-cycle reactions in its core (e.g. [70]).

However, because this process is so sub-dominant in solar-like stars, it was never actually seen in the Sun before the measurement presented in section 3.4. Furthermore, because the reaction rate of CNO is very dependent on the abundances of carbon and nitrogen (which act as catalysts) in the Sun, a measurement of the flux of CNO neutrinos can be used to infer properties of the chemical composition of the core of the Sun, as well as to provide a fundamental test of the standard solar model. Because of a radiative zone in which no convection can occur, the core of the Sun is thought to be chemically

decoupled from the surface, therefore making the flux of CNO neutrinos a probe into the initial chemical composition of the Sun at its genesis. A precision measurement even has the potential to solve one of the main open questions in the field of solar physics, the solar metallicity problem (see section 1.3.1).

1.4 Reactor neutrinos

Reactor neutrinos are of great significance in neutrino physics, having been the source of the first ever detected neutrinos in 1956, see above. They are today still a great source of neutrino physics, with today's world being much more reliant on nuclear power than in the 50s. The main source of reactor neutrinos (over 99%) are beta decays along the fission chains of four isotopes: ^{235}U, ^{238}U, ^{239}Pu, and ^{241}Pu, with an average six anti-electron-neutrinos per fission of about 1.5 MeV each (on average over all neutrinos). The maximum energy of the neutrinos is roughly 10 MeV. However, only an average of two out of the six neutrinos have energies above the 1.8 MeV threshold of IBD detection (see section 1.6.5) [71]. One can calculate expected neutrino fluxes from a reactor using [72]:

$$\Phi(E_\nu) = \frac{P_{th}}{\sum_i^{isotopes} f_i E_i} \sum_i^{isotopes} f_i \Phi_i(E_\nu), \tag{9}$$

where P_{th} is the reactor's thermal power, f_i is the fission fraction for the isotope i in the reactor, E_i is the energy released per fission (about 200 MeV per isotope on average), and $\Phi_i(E_\nu)$ is that isotope's neutrino spectrum. Reactor neutrinos are used to study neutrinos themselves, holding the leading results on θ_{13} from Daya Bay [73] and on Δm^2_{21} from KamLand [74]. They also constitute an important factor in the search for sterile neutrinos because of the extremely short baselines that they can measure. See fig. 12 for an example of a typical reactor neutrino spectrum. There are currently two main so-called *anomalies* in the field which look like they could teach us something about nuclear reactors, although more data is needed. Notably, the JUNO experiment (see fig. 13) will use reactor neutrinos to determine the MH (see section 1.7) with a 3σ significance.

The first anomaly, commonly referred to as the *reactor-antineutrino-anomaly* (RAA) is characterised by the fact that very short baseline reactor neutrino experiments measured a lower absolute antineutrino rate than was expected from models. It started when the models were refined in 2011 [75], which revealed a discrepancy of 6%. This was later confirmed by several independent experiments [39, 73, 76]. While a sterile neutrino hypothesis was the first suggestion to solve this problem [72], later Daya Bay [77] and Reno [78] independently showed that the IBD yield of ^{235}U was higher than expected at the 3σ level, and [79] further showed that updated measurements involving ^{235}U

and ^{239}Pu decays could also lead to a solution that doesn't involve sterile neutrinos, though more data is needed. Additionally, very short baseline experiments like STEREO excluded the sterile neutrino parameter space of the RAA at the 99.9% C.L. [80].

The second anomaly is the so-called *reactor shape anomaly*, referring measurements of an unexpected bump in the antineutrino rate near 5 MeV. This was measured by several experiments [81–83] including Daya Bay [84], which all showed that the significance of the bump was correlated (with a significance of up to 4σ [85]) to the thermal power of the reactor. This of course again strongly suggests that the origin of the anomaly lies in our understanding of nuclear reactor cores and not in neutrino properties beyond the three-neutrino framework. JUNO-TAO, see section 2.2.2, is expected to be a main contributor in the solution of both anomalies thanks to its unprecedentedly precise flux measurement.

1.5 Other neutrino sources

Atmospheric neutrinos are created in the atmosphere in the cosmic air-showers originating from the collision of high-energy cosmic-rays in our atmosphere. Most cosmic-rays are protons and the majority of the particles created in the showers are pions (and kaons). The pions quickly decay, sometimes followed by a muon decay:

$$\begin{aligned} \pi &\to \mu + \nu_\mu, \\ \mu &\to e + \nu_e + \nu_\mu. \end{aligned} \tag{10}$$

As one can see, neutrinos are created in all of those decays. From simple considerations using the branching ratios of the different decays, one can see that the ratio of the total number of muon flavour neutrinos to electron flavour ones is around two. This number increases at energies higher than about 2 GeV because the muons start to reach Earth before decaying. Atmospheric neutrinos have been used to provide the first evidence for neutrino oscillations by SuperKamiokande [12] which has led to the first measurement of θ_{23} and of Δm^2_{32}. Other atmospheric neutrino measurements have been performed by the Lake Baikal experiment [86], IceCube [87], Km3Net [88], and ANTARES [89]. These have contributed to oscillation parameter measurements, and to the observation, by SuperKamiokande [90] and DeepCore [91], of tau neutrinos, which are not created in the atmosphere in significant amounts but result from oscillations. Because some of the neutrinos reach detectors directly and some reach them after having travelled through Earth, atmospheric neutrinos can be measured with baselines ranging from 10 km to 1×10^4 km (with energies from 0.1 GeV to 1×10^5 GeV). This also makes atmospheric neutrino measurements sensitive to the MSW effect, and by proxy to the

neutrino mass hierarchy (MH), as will be exploited in the JUNO experiments to provide some additional sensitivity [92]. PINGU [93] (whose data is complementary to JUNO's [94]), HyperKamiokande [95], and ORCA [96] will also be sensitive to the MH and can be expected to provide precision θ_{23} and of Δm^2_{32} measurements.

Supernova neutrinos are created in huge bursts (of around 10^{58}) in a Supernova. These emissions occurs in three phases: *Burst*, *Accretion*, and *Cooling* [97–99]. *Burst* neutrinos are emitted mostly by electron capture on protons, stemming from the shock wave from the core-collapse ramming into in-falling high-Z matter. They are emitted on time-scales of a few tens of microseconds after the collapse with predominantly ν_e. *Accretion* neutrinos are, as the name suggests, emitted during a period of stagnation of the shock wave where heat is generated through accretion flow. The heat is mostly "evacuated" through thermal neutrino emission of all flavours (ν and $\bar{\nu}$) and through charged-current electron conversion. Because the $\mathcal{O}$(MeV) energies are too low to generate heavier leptons, most neutrinos emitted here are also electron flavoured, though the other leptonic counterparts are also generated, generally in the deeper, hotter regions. These processes produce a region of net energy gain behind the shock wave, which builds up pressure and re-accelerates the explosion until the accretion flow is terminated [92]. This happens on the time scale of 1 s. Finally, over $\mathcal{O}(10\,\mathrm{s})$, *Cooling* neutrinos are emitted during the so-called deleptonisation of the gas surrounding the proto-neutron star.

They have so far only been measured once in 1987 when SuperKamiokande measured eleven neutrinos in 13 s [100], Baksan five in 9 s [101], and Brookhaven eight in 6 s [102]. The events were detected with average energies slightly higher than 10 MeV, mostly through inverse beta decays (see section 1.6.5). This last measurement already revolutionised our understanding of supernovæ and since then, the world has been preparing for the next supernova with the so-called SNEWS [103] system. It consists of several neutrino detectors that, in the case of a supernova, are expected to measure huge signals hours before any electromagnetic radiation can reach Earth. This will give time to other observatories to orient their dishes and prepare. A future measurement will not only immensely contribute to our understanding of supernovæ, but also provide information about neutrinos themselves, as *e.g.* JUNO is in that case expected to set a $m_\nu < (0.83 \pm 0.24)\,\mathrm{eV}$ (95% CL) upper limit on the absolute neutrino mass [104]. It is however unclear when the next supernova will occur as only one to three are expected to occur per century in our galaxy. However, about one supernova occurs every second in the universe, leading to a theorised diffuse supernova neutrino background with a very low rate. This flux has never been observed, but is expected to be measured with

a 20-33% uncertainty in the next few decades by a combination of the data from HyperKamiokande, DUNE, and JUNO [105], though JUNO alone is expected to reach a discovery power of 3σ in ten years [106].

Geo-neutrinos from primordial nuclei dispersed inside the Earth's crust and mantle [107] can be measured with liquid scintillator detectors. The two non-negligible isotopes that can be measured (above the 1.8 MeV IBD threshold, see section 1.6.5) are from ^{238}U and ^{232}Th. Their spectra reaches from that threshold to 3.27 MeV and 2.25 MeV, respectively. Geo-neutrinos are extremely useful particles in the context of geology as they can give direct information about the amount, distribution, and relative quantities of primordial ^{238}U and ^{232}Th, which is otherwise impossible to measure directly below just a few kilometres. More specifically, geo-neutrinos can answer questions about the radiogenic heat production (currently known only with high uncertainties in the mantle) and about the U/Th ratio, which is an important parameter in models of Earth's formation [108]. KamLand and Borexino are the only experiments to have measured geo-neutrinos to date with about 169 in six years [109] and 53 in nine [110], respectively. Future high-statistic experiments, such as JUNO which is expected to measure about 400 geo-neutrinos per year [111], will be able to additionally measure the U/Th ratio, which is currently not feasible because of low statistics. SNO+ is also expected to achieve a high signal-to-noise ratio, mostly because of the local abscence of reactor neutrinos [112]. Future experiments are also planned, namely a 3 kt detector at Jinping expected to reach 25% [113] precision on the U/Th ratio, and a 10 kt liquid scintillator at Baksan expected to reach 10% [114].

Accelerator neutrinos have been used since the 1960s [115, 116], to for example prove the existence of more than one kind of neutrino [6]. Accelerator neutrinos are also promising sources to study the MH and to try and measure δ_{CP}. In case of maximal CP violation, the T2K-II is for example expected to reach a 3σ sensitivity to δ_{CP}, and NOvA will be sensitive to the mass hierarchy [72]. When combining data from T2K, NOvA, and JUNO, one can reach conclusive evidence for the MH for all values of δ_{CP} [117]. In the future, DUNE and Hyper-K can also be expected to measure the MH for all values of δ_{CP} and to give a high sensitivity to CP violation [118]. Current neutrino beams are generated by colliding protons onto a target, which produces many pions and kaons. These charged hadrons are directed in a specific direction and filtered by electric charge before decaying into leptons and their corresponding neutrinos. This has been used to create neutrinos with energies from 0.5 GeV to 500 GeV [119, 120]. By utilising a low-energy pion beam and by emitting neutrinos under a certain laboratory angle to said beam, it is possible to use the so-called decay-at-rest method to get a

neutrino beam with a sharp energy distribution [121]. Using this method is currently under investigation with kaons as well [122]. In the future, other kinds of neutrino beams are expected to be made, for example using muon storage rings [123].

1.6 Liquid scintillator neutrino detection

1.6.1 Liquid scintillator detection principle

The principle of particle detection in liquid scintillator experiments is tied to the behaviour of different molecules inside it. Liquid scintillator is based on an organic solvent whose molecules can be excited by the passage of charged particles, the de-excitation of those molecules then produces light in the UV or visible spectrum in a process called luminescence [124]. Typical solvents for liquid scintillator are linear-alkylbenzene (LAB) as will be used in JUNO and 1,2,4-trimethylbenzene (pseudocumene) as was used in Borexino. Because the emitted light has a high probability of being re-absorbed by neighbouring molecules, which would prevent light from reaching any detector, one must add a so-called wavelength shifter to the liquid which absorbs the photons before re-emitting them with a longer wavelength. Even in an ideal case, charged particles deposit energy in the liquid, which is then to the first order linearly converted to visible light, but with deviations especially at low energies, see section 1.6.2. That light can subsequently be detected, for example using PMTs.

1.6.2 Charged particles in liquid scintillator

Energy losses As described by the Bethe-Bloch formula [20], which includes the processes of ionisation and excitation, heavy charged particles loose energy when passing through matter. This formula is in general valid in the range of $0.1 < \beta\gamma < 1000$, but for lighter particles, such as electrons, correction terms need to be included to consider Bremsstrahlung. About three percent of the ionisation energy ends up as scintillation light [20].

Non-linearity Along the tracks of low-mass particles, the energy is usually deposited rather continuously. At the end of the particle tracks however, the density of excited molecules gets so high that neighbouring excited molecules start interacting with each other, which reduces the total amount of produced light. This effect is called quenching and was empirically parametrised by Birk [124]:

$$\langle \frac{dL}{dX} \rangle = L_0 \frac{\langle \frac{dE}{dX} \rangle}{1 + kB \langle \frac{dE}{dX} \rangle}, \tag{11}$$

where L is the light yield, L_0 a normalisation constant, E is the deposited energy, X is the *depth* (distance over mass density), and kB is a material-dependent constant [125]. The effective light yield loss by quenching affects alpha particles about ten times more than electrons.

Cherenkov light The well-known Cherenkov effect occurs when a charged particle moves faster than electromagnetic waves in a given medium. While in liquid scintillator the amount of energy lost through Cherenkov radiation is small compared to scintillation, it is nevertheless an important factor to consider, and it causes additional non-linearity in the light yield. It can be particularly useful because, unlike scintillation light which is emitted isotropically, it carries directional information about the incident particle. This was recently shown to be useful even in liquid scintillator detectors in the context of solar neutrino detection (through elastic scattering, see section 1.6.5) by using the so-called *Correlated and Integrated Directionality* [126, 127]. Next-generation detectors will probably use liquid scintillator diluted in water in efforts to increase the amount of Cherenkov light - while keeping the light yield and low threshold advantages from the liquid scintillator technique [128]. The wave front of Cherenkov light is emitted at an angle of $\cos\theta = 1/n\beta$, where θ is the Cherenkov angle, n is the (wavelength dependent) refractive index, and β is the speed of the particle divided by c. For relativistic electrons in liquid scintillator, which has a refractive index of about 1.5 at wavelengths of 235 nm which is typical of scintillation light, the Cherenkov radiation will be emitted at a maximum angle of 48° but only for electrons with energies above about 256 keV.

Scintillation time profile Because the de-excitation of the scintillating atoms is not instantaneous, the scintillation light generated by one particle is generally emitted along a time profile which can be parametrised by the sum of (about 3-4) exponential distributions as in:

$$f(t) = \sum_i \frac{q_i}{\tau_i} e^{-t/\tau_i}, \tag{12}$$

where t is the time, q_i and τ_i are the exponential parameters. Representative values for JUNO can be seen in table 2. This is very useful because different particles usually lead to different time profiles [129], which can be used to discriminate between them. As one can see from the table, there are for example significant differences between electrons and alpha particles, especially in the medium to long components.

particle type	e^-/e^+		α	
i	q_i	τ_i (ns)	q_i	τ_i (ns)
1	0.799	4.93	0.65	4.93
2	0.171	20.6	0.228	35
3	0.03	190	0.123	220

TABLE 2: Representative parameter values for liquid scintillator time profiles for electrons and positrons in JUNO, as well as alpha particles [92, 130].

1.6.3 Gammas and neutrons

Neutral particles can also end up indirectly depositing *visible energy* (referring to the energy that is converted in detectable light) in liquid scintillator via their interactions with other particles. Gammas with energies above $2m_e$ will for example produce electron-positron pairs. Below that energy, gammas can interact through three interactions, two of which generate light in the detector. The first interaction is *Rayleigh* scattering: a non-light-producing coherent scattering which transfers momentum but no energy. The other two are *Compton* scattering and the photo-effect. The former is an incoherent scattering on an electron in which part of the energy of the gamma is transferred to the electron, typically ionising its parent atom (and thus emitting light!). The latter usually occurs only at very low energies (dominant below about 20 keV for LS, see fig. 7), after the gamma has undergone several Compton (and possible Rayleigh) scatterings. In this process the gamma is fully absorbed by the electron. For typical MeV energies Compton scattering is the dominant process, and the relatively large mean free path of gammas in the LS means that the energy deposition is far from point-like. For events that are close enough to the edge of the detector, this can cause some of the energy to be deposited outside of the LS - thereby also creating non-linearity.

Thermal neutrons are typically captured on free protons in the liquid scintillator [131]. This forms a deuteron and emits a gamma with an energy of 2.2 MeV (the deuteron's binding energy) after an average time of 200 µs [92]. There is however also a small probability for the neutron to be captured by a carbon atom, which similarly emits a gamma with an energy of 4.95 MeV. Higher energy neutrons (often called *fast neutrons*) will typically transfer part of its energy to protons through scattering (the protons then cause scintillation) until their energies are low enough for them to be absorbed on protons as well. The TAO detector (see section 2.2.2) uses a special Gadolinium-doped scintillator with a high probability of capturing those neutrons instead and thereby producing 8 MeV gammas after on average 30 µs (instead of 200 µs) [132]. This technique, which was also used in other reactor neutrino experiments [38, 82, 133], has the advantage

[2] http://nucleardata.nuclear.lu.se/toi/nuclide.asp?iZA=300065

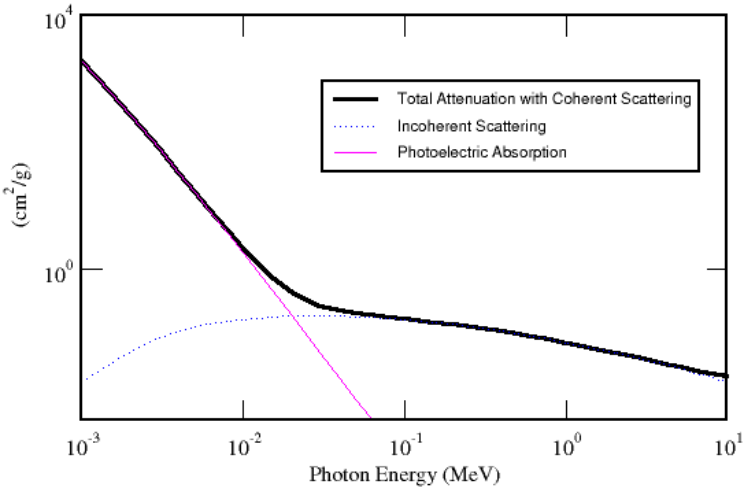


FIGURE 7: Interaction cross sections in LAB (see section 1.6) as a function of gamma energy. As one can see, the photo-effect remains the dominant interaction process until about 20 keV, when Compton scattering quickly takes over. Image credit: Philipp Kampmann through NIST XCOM database[3]

of a larger cross section - as well as a higher efficiency and background rejection in the neutron tagging thanks to the higher gamma energy.

1.6.4 Positrons

Because of the prevalence of electrons in matter, positrons in scintillator will generally annihilate after having deposited their kinetic energy as described above, thereby releasing two gammas with an energy of $m_e = 511\,\text{keV}$ each - which can be detected thanks to the processes mentioned above. There is however also a probability that the positron will first form an exotic positronium atom, which will either be in the form of ortho- or of para-positronium, depending on the spin of the system. Because the decay of ortho-positronium (spin 1) into two gammas is forbidden, it has a much longer lifetime than its counterpart [134–136]. This results in a multi-modal and on average longer scintillation time profile for positrons when compared to electrons. Because of this and the additional difference in time/spacial behaviour between positron and electron events due to the back-to-back gamma annihilation, it is possible to differentiate both particle types in measurements [130, 137].

1.6.5 Neutrino interactions in liquid scintillator

In liquid scintillator detectors, two neutrino interactions with matter are mainly used. The first one, Inverse Beta Decay (IBD), occurs through the charged current interaction of an antineutrino on a proton, resulting in a neutron and a positron, see fig. 8(a).

At $\mathcal{O}$(MeV), this is the only reaction through which electron antineutrinos have been measured in LS detectors. This is for example used in JUNO to very efficiently detect reactor antineutrinos. This offers a very strong coincidence signal because the positron quickly (few nanoseconds) annihilates, thereby creating two very visible gammas with $E_\gamma = m_e$, and then when the neutron is captured, one detects a 2.2 MeV or 4.9 MeV gamma signal, see above, after on average about 250 µs [92]. The space coincidence due to the low mean-free-path of the neutron can thereby also be used. The energy carried off by the prompt positron is $E_{e^+} = E_{\bar{\nu}_e} - 1.806\,\text{MeV}$ because of the rest-mass difference between protons and neutrons. It then deposits a total *visible energy* of $E_{vis} = E_{\bar{\nu}_e} - 0.784\,\text{MeV}$ in the scintillator because of the two 511 keV gammas from the annihilation. The drawbacks of this method are the energy threshold, of 1.806 MeV, and that it can only be used to detect electron flavoured antineutrinos.

The second kind of interaction is elastic scattering (ES) off electrons. This can happen either through W (CC) or Z (NC) boson exchange. The former can again only occur with electron neutrinos, while the latter can occur with all flavours, which is why the cross-section of this interaction is (about six times) higher for electron neutrinos than for the others. The interactions can be seen in figs. 8(b) and 8(c). Elastic scattering signals transfer a lot of energy to the electron, whose so-obtained kinetic energy is then visible as a flash of light in the detector, see above. For electron antineutrinos, the cross-section of ES is of course dominant below the IBD threshold, but otherwise it is 1-2 orders of magnitude higher for IBD in the typical energy ranges [138]. Further, while the signals from IBD have sharp signatures through coincidence, it is more difficult to detect ES signals which only appear as single flashes of light. Further, because of the kinematics behind the electron recoil spectra following an ES event, even mono-energetic neutrinos will produce a relatively flat recoil energy spectrum with a Compton-like edge at the initial neutrino energy.

1.7 Open questions in neutrino physics

This chapter has reviewed and shown many known properties of neutrinos and their sources and has given brief insights into detectors and methods that have led to this knowledge. Neutrino physics is however still a relatively young field with a multitude of open questions and possibilities. These questions are all the more interesting with essentially most of neutrino physics being physics beyond the Standard Model. This section briefly covers the most important open questions in neutrino physics and how they might be answered.

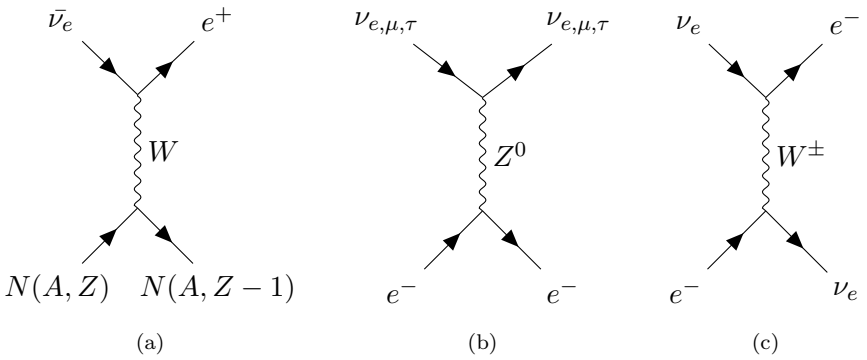


FIGURE 8: Neutrino interactions in liquid scintillator. (a) Inverse beta decay in which only electron-flavour neutrinos can participate at $\mathcal{O}$(MeV) energies. (b) and (c) show elastic scattering off an electron, once for neutral current (all flavours) and charged current (only electron flavour again) interactions (note that the equivalent diagrams exist for an incoming anti-neutrino).

Mass hierarchy The MH, see section 1.2, is an important parameter not only in particle physics but also in astrophysics and cosmology [92]. It is for example important in the search for neutrinoless double beta decay [139], see below, and affects supernova neutrinos (through matter effects) [140] as well as many theoretical models [141–143]. As will be seen below and is shown in fig. 9, the MH significantly affects the allowed parameter space for $m_{\nu_e}^{\text{eff}}$ for low neutrino masses. It also has a big influence on the search for δ_{CP} [35]. The JUNO experiment will play a large role in determining the MH, see section 2.2.1, as it is expected to reach 3σ in six years. Because it doesn't rely on matter effects for its determination, JUNO's data offers a lot of synergy with other experiments, and could *e.g.* through a joint effort with PINGU reach a 5σ determination in a few years [94]. On JUNO's side, this sensitivity comes from the interplay between two different oscillations (Δm^2_{32} and Δm^2_{31}, see section 2.2.1), and on PINGU's side, this mainly comes from the study of atmospheric neutrinos through matter effects. Other experiments, such as the future DUNE experiment [144], are sensitive to the MH through long baseline electron (anti)neutrino appearance in muon (anti)neutrino beams. Overall, we can expect this question to be answered in about the next decade.

Mass origin There are currently two preferred models that could explain the neutrino mass while preserving gauge invariance [20]. The first one is of so-called *Dirac neutrinos*. This involves adding a mass term to the Standard Model Lagrangian which in the particular case of three sterile neutrinos can be identified with the right-handed component of a four-spinor neutrino field. This scenario can however not explain why neutrinos are much lighter than their corresponding charged fermions. The second model, essentially an extension of the first, is of so-called *Majorana neutrinos* in which one field describes both the neutrino and anti-neutrino states. The Majorana term is proportional to a

mass term M_N. In the specific case where M_N's eigenstates are much higher than the mass scale of electroweak symmetry breaking, then the so-called *see-saw* mechanism, which involves three light neutrinos and three heavy neutrinos, could explain the origin of neutrino masses. In this mechanism, the masses of the light neutrinos are proportional to M_N^{-1}, those of the heavy ones are proportional to M_N, hence the name *see-saw*. This would explain the difference in the neutrino/lepton masses and is expected in SO(10) GUTs [145, 146]. Should this be the case, one would expect the detection of neutrinoless double beta decay to be possible, see below.

Absolute mass scale Within the normal hierarchy, one can approximate that m_2 is probably not far from $\sqrt{\Delta m_{21}^2} \approx 8.6 \times 10^{-3}\,\text{eV}$, and that m_3 can be estimated with $\sqrt{\Delta m_{32}^2 + \Delta m_{21}^2} \approx 0.05\,\text{eV}$. Similar approximations within the inverted hierarchies lead to about 0.05 eV for both m_1 and m_2 [20]. There exist different approaches which are currently trying to reach those levels in order to determine the absolute mass scale. Perhaps the most notable example uses an end-point measurement of beta decays to get a direct neutrino-mass probe from kinematics. The Katrin experiment, using the kinematics of tritium decay and detailed analysis of the beta-spectrum endpoint, was recently able to set an upper limit of 0.8 eV (90% C.L.) [3] on the effective electron mass term $m_{\nu_e}^{\text{eff}}$:

$$\begin{aligned} m_{\nu_e}^{\text{eff}} &= \sqrt{\sum_i m_i^2 |\mathcal{U}_{ei}|} \\ &= \begin{cases} \sqrt{m_0^2 + \Delta m_{21}^2 (1 - c_{13}^2 c_{12}^2) + \Delta m_{32}^2 s_{13}^2} & \text{, normal hierarchy} \\ \sqrt{m_0^2 + \Delta m_{21}^2 c_{13}^2 c_{12}^2 - \Delta m_{32}^2 s_{13}^2} & \text{, inverted hierarchy} \end{cases} \end{aligned} \tag{13}$$

where m_0 is the lightest neutrino mass: $m_1 (m_3)$ for the normal(inverted) hierarchy [20]. Equation (13) is valid for all beta decays, but Katrin's measurement was achieved by injecting molecular tritium into a series of superconducting magnets leading to a large MAC-E-filter which acts as a high-pass filter. Thanks to a four-orders-of-magnitude reduction in the strength of the magnetic field from the entrance of the spectrometer to its central plane, a narrow filter with a width of 2.8 eV could be created, leading to a precise measurement of the beta-electron spectrum endpoint. The interpretation of this result in the context of the three mass terms (m_1, m_2, m_3) can be understood more directly by looking at fig. 9. As one can also see from eq. (13), a constraint on $m_{\nu_e}^{\text{eff}}$ has a different influence on m_0, as reflected in fig. 9.

As one can see, there are already strong experimental bounds on the relationship between the effective neutrino mass and the mass of the lightest neutrino state. The important

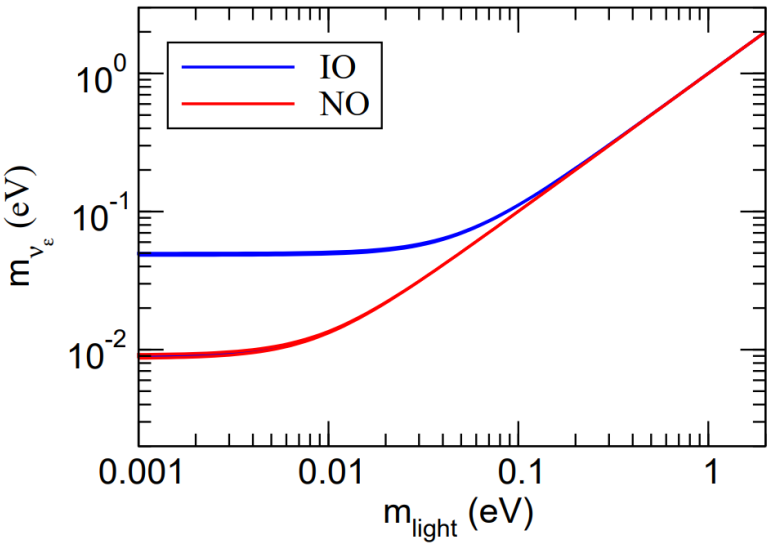


FIGURE 9: Allowed 95% ranges for $m_{\nu_e}^{\text{eff}}$ (see text) as a function of the lightest neutrino mass. One can see that if the lightest neutrino mass is below about 0.1 eV, then one could start probing the normal and inverted hierarchies (NO and IO in the plot, respectively). The tightest upper limit on the lightest neutrino mass, from cosmological measurements, is 0.09 eV (95% C.L.) [151]. This figure was taken from [20].

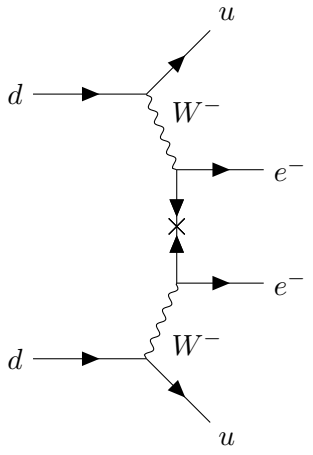


FIGURE 10: Neutrinoless double beta decay (both incoming down quarks belong to respective neutrons): the electron neutrino in the center is a Majorana particle.

thing about Katrin's current result is that it is limited by statistical uncertainties. Overall, the collaboration expects to reach an upper limit of 0.2 eV, which is very close to the aforementioned lower bounds from the oscillation parameters. Further efforts are also ongoing, including Project-8[4] which plans to use cyclotron radiation to possibly reach a sensitivity of 40 meV [147, 148] and *e.g.* the Holmes experiment [149] which will be somewhat similar to Katrin but using the ^{163}Ho decay (which has a smaller Q value) instead of tritium [150], which is currently investigating how well it can scale up.

[4] https://www.project8.org/

If neutrinos are indeed Majorana particles, it is also possible to measure the absolute mass scale through a measurement of *neutrinoless double beta decay* ($0\nu\beta\beta$) [152, 153]. This is a rare decay branch where two close-by neutrons simultaneously undergo beta decay. Instead of each emitting a neutrino and an electron, the Majorana nature of the neutrino would instead cause only the electrons to be emitted, see fig. 10. Getting a measurement of the rate of $0\nu\beta\beta$ would constitute a measurement of $|m_{ee}| = \sum_i m_i \mathcal{U}_{ei}^2$. The current best upper limits on $|m_{ee}|$ reach levels of 60 meV [154, 155].

Finally, the best upper limits on the lightest neutrino mass are from cosmology, where by using cosmic microwave background temperature fluctuation and polarisation data combined with supernovæ Ia luminosity distances, data from galaxy clustering, and large scale spectroscopic surveys, one reaches a value of 0.09 eV (95% C.L.). While this result is highly model dependent, it is by far the strongest constraint on the lightest neutrino mass and the only field where future experiments are guaranteed to reach a neutrino mass measurement, should planned telescopes come online.

CP violation Given the lack of understanding in the face of our universe's baryon asymmetry, CP violation is of course a very important field of research, and could be explained by the lepton sector through the δ_{CP} phase. CP violating processes through δ_{CP} can only occur if all PMNS mixing angles are non-zero and if δ_{CP} is different than 0° and 180° [35]. While it is already known that all mixing angles are non-zero, the value of δ_{CP} is currently under investigation. The value of δ_{CP} is of great importance also in the determination of the MH with accelerator experiments, as explained in section 1.5 under "**Accelerator neutrinos**". For example, NOvA will be able to determine the MH with $4-5\sigma$ in case of maximal violation, but with less than 1σ in case of no violation - though these results also depend of the θ_{23} octant [72]. In the future, DUNE and Hyper-K will reach a $> 5\sigma$ sensitivity to the MH for all values of δ_{CP} [118].

θ_{23} **Octant** Results of global fit analyses in the three neutrino framework show a preference for a best-fit value of $\sin^2 2\theta_{23} = 0.573^{+0.018}_{-0.023}$ [27]. However, this is a very degenerate solution which could just as easily be closer to 0.45 - with only a slight preference for the quoted value. This degeneracy is in fact a main limiting factor for the sensitivity of neutrino beam experiments to θ_{13} [27]. It can however be measured by large baseline beam experiments through the appearance(disappearance) of $\nu_e(\nu_\mu)$ which has a direct dependence on $\sin^2 2\theta_{23}$, though these results in general depend on δ_{CP} and on the MH. It can also be measured through sub-leading modes of atmospheric neutrino oscillations [20]. This is something that the JUNO experiment will be able to shed some light on [156].

Sterile neutrinos While the three-neutrino framework is in general a very successful theory, several experiments have reported different kinds of anomalies (though none at the 5σ level). In 2001, the LSND experiment found a 3.8σ excess of $\bar{\nu}_e$ appearance (from a ν_μ beam) [157]. MiniBoone reported a similar 4.8σ excess in both $\bar{\nu}_e$ and ν_e appearance [158]. Gallium experiments (GALLEX and SAGE) meanwhile found a 2.5σ deficit in the expected neutrino rates from a calibration-related radioactive source [159]. This was recently raised to the level of 4σ by BEST [160], which confirmed the deficit. There are other possible hints from reactor experiments, though these seem to be more readily explainable by reactor physics. When taking the electron (anti)neutrino appearance signals and the electron (anti)neutrino disappearance signals on their own, one find in both cases that they are explainable under a so-called (3+1) framework. This is similar to the three-neutrino paradigm but with an additional *sterile* neutrino with a mass square difference of $\Delta m^2_{41} \approx \mathcal{O}(\text{eV}^2)$ which can participate in oscillations through a non-zero angle θ_{14}, but does not interact with any bosons, not even in weak interactions. However this model, under the assumption of unitarity, necessarily implies the existence of muon (anti)neutrino disappearance, which has not been observed. The existing limits on this disappearance are in stark tension with the aforementioned eV-scale results [161], especially when considering recent IceCube atmospheric neutrino data [162].

Because of the strong tensions and lack of unified understanding, many experiments are of course ongoing and planned to try and reject this hypothesis. For example, SBND and Icarus, both planned to be installed on the Fermilab neutrino beam, are designed to, together with MicroBoone [163], explore the phase-space suggested by LSND and MiniBoone [72, 164]. Reactor experiments also play a part. Notably Neutrino-4 has reported a 2.7σ signal [165], though it is in tension with results from Daya Bay, Bugey-3, and MINOS+ [166], as well as with PROSPECT data [167]. However, the parameters inferred from BEST do agree with Neutrino-4's results. These results are already planned to be tested in the next few years by DANSS [168] and PROSPECT [169] after detector upgrades. There are of course many more efforts which, for the sake of brevity, are not covered here, see [170] for a more detailed review.

Non-standard interactions A possibility of new physics in the realm of neutrinos arises from adding vector-like couplings between neutrinos and other fermions to the SM Lagrangian [72]. Relatively simple models can reasonably be expected to lead to coherent scattering in matter, similar to the MSW effect for electron neutrinos, which could be measurable in long baseline beam experiments [171]. It could also allow to use neutrinos to probe possible heavy new particles. When including up and down quarks, such new physics would influence the cross-sections of coherent elastic neutrino-nucleus scattering (CEνNS) [172, 173]. One very significant consequence of such new physics is that it could

mimic a signal for leptonic CP-violation or lead to a false determination of the MH [172], which could however be cross-checked by precision PMNS-matrix measurements, such as the ones that will be performed by JUNO.

Magnetic moment The observation of a neutrino magnetic moment would necessarily also lead to new physics [72]. Such an effect would be for example a consequence of neutrinos being Dirac particles, which would create a magnetic moment through charged W couplings, albeit a small one. In fact, a straightforward approach leads to a prediction for μ_ν of about $3.2 \cdot 10^{-19}(m_\nu/\mathrm{eV})\mu_B$, where μ_B is the Bohr magneton. Because the cross-section of neutrino elastic scattering on electrons is dependent on μ_ν, upper limits were already set but they only reach levels of $\mathcal{O}(10^{-11})\mu_B$, which is orders of magnitude away [174]. Astrophysical sources may in the future reach more sensitivity, as for example the neutrino measurements related to SN1987A already led to a limit of $\mathcal{O}(10^{-12})\mu_B$ [175].

2 Experiments

2.1 Borexino

Borexino was initially built to detect ^{7}Be solar neutrinos (see section 1.3.2), which it achieved within only a few months of data taking [176]. As discussed below, it quickly went on to substantially surpass these initial expectations and continued making groundbreaking discoveries until the very end of its data-taking, for example when in 2020 it was used to achieve the first discovery of CNO neutrinos [68], which constitutes a large research topic of this thesis (see section 3.4).

2.1.1 Physics goals

Thanks to more than a decade of careful preparations and developments, Borexino achieved its initial goal of measuring ^{7}Be solar neutrinos only months after the start of data taking in 2007 [176]. This was made possible by the unprecedented radiopurity of its materials, which is the pillar on which of all it physics achievements rests. Only a few years later, Borexino scientists were able to measure the rate of ^{8}B neutrinos with an unprecedentedly low energy threshold of 3 MeV [177], which was made possible through the hitherto mostly unexplored technology of liquid scintillator detection. Previous analyses with water-Cherenkov detectors used a 5 MeV threshold [178]. Shortly after, having accumulated more statistics and with the question of proving the existence of ^{7}Be neutrinos out of the way, Borexino scientists published a precision measurement of the rate of ^{7}Be neutrinos (at a level of 4.8%) [179] and measured *pep* neutrinos for the first time [180, 181]. Borexino also became one of only two experiments to ever measure geo-neutrinos [182], paving the way for an exciting new field.

Borexino then underwent a large purification campaign which marked the beginning of the so-called *Phase-II* dataset, which encompasses data from December 2011 to May 2016. The campaign led to a reduction in the rates of ^{85}Kr from about 30 cpd/100t to 6.8 cpd/100t (counts per day per 100 tonnes of scintillator), ^{210}Bi from about 41 cpd/100t to 17.5 cpd/100t, and ^{210}Po from over 500 cpd/100t to 260 cpd/100t. Additionally, the contamination levels of ^{238}U and ^{232}Th were reported to be less than $9.4 \times 10^{-20}\,\mathrm{g\,g^{-1}}$ and $5.7 \times 10^{-19}\,\mathrm{g\,g^{-1}}$ at the 95% C.L., respectively. This led to the first measurement of *pp* neutrinos [183] and culminated in a comprehensive measurement of neutrinos from the entire pp-chain [46]. The final results of Phase-II on the solar neutrino fluxes can be seen in table 3. Using those results in a calculation of the electron neutrino survival probability from the Sun, a purely vacuum hypothesis (no matter effects) could be rejected at a 98.2% C.L., and when using the ^{7}Be and ^{8}B neutrino fluxes - which are

most sensitive (except CNO) to changes in metallicity, the low-metallicity model could be disfavoured at the 96.6% C.L.

Solar neutrino	Flux ($\mathrm{cm^{-2}\,s^{-1}}$)
pp	$(6.1 \pm 0.5^{+0.3}_{-0.5}) \cdot 10^{10}$
^{7}Be	$(4.99 \pm 0.11^{+0.06}_{-0.08}) \cdot 10^{9}$
pep (HZ)	$(1.27 \pm 0.19^{+0.08}_{-0.12}) \cdot 10^{8}$
pep (LZ)	$(1.39 \pm 0.19^{+0.08}_{-0.13}) \cdot 10^{8}$
^{8}B	$(5.68^{+0.39+0.03}_{-0.41-0.03}) \cdot 10^{6}$

TABLE 3: Solar neutrino fluxes as measured by Borexino in the Phase-II dataset [46]. The values are given with statistical and systematic uncertainties, respectively. *pep* results are shown under the assumption of a High- and of a Low-Metallicity SSM (see section 1.3.1).

At the end of 2015, a thermal insulation of Borexino was completed, followed by the installation of an active temperature control system, in early 2016, which was installed on top of the detector [184], see section 2.1.2. This stabilised the temperature gradient inside of Borexino, which was previously known to fluctuate seasonally. After the thermal insulation, in the new *Phase-III* dataset spanning from July 2016 to February 2020, a very stable temperature gradient of $0.5\,^{\circ}\,\mathrm{m^{-1}}$ was established inside of the detector, between the cold rock ($\approx 8^{\circ}$) below and the air ($\approx 15^{\circ}$) above. This stable gradient served to counteract the effects of gravity to minimise convection in the fiducial volume. As a consequence the rates of polonium inside of the detector were stabilised (see section 3.5). This made it possible to constrain the rate of ^{210}Bi, a main background in the CNO measurement of Borexino, and eventually led to the first experimental evidence of CNO neutrinos [68], see section 3.7.

With a varied physics program, Borexino also *e.g.* provided leading limits on a neutrino magnetic moment [185], measured modulations of the cosmic-muon flux at Gran Sasso [186], and set constraints on flavour-diagonal non-standard neutrino interactions [187]. Borexino can also be used for astrophysical searches, for example in [188] where it set the strongest upper limits on FRB-associated neutrino fluences of all flavours in the 0.5 MeV to 50 MeV neutrino energy range. More recently, Borexino was used to prove that even in large liquid scintillator detectors, the small fraction of Cherenkov light that precedes the slower but dominant scintillation light could be used to measure the rate of solar neutrinos [126, 127], thus paving the way for future large liquid scintillator detectors (like JUNO) to use directional information as background suppression and as disentanglement of signals.

2.1.2 Detector design and status

The Borexino detector, see fig. 11, consists of onion-like concentric shells. The innermost volume, inside of the *inner nylon vessel* (IV), contains the liquid scintillator: the target of the experiment. Borexino's scintillator, that was the most radio-pure neutrino detector material in the world, consisted of nominally 280 t of a 1,2,4-trimethylbenzene (pseudocumene) solvent with 2,5-diphenyloxazole (PPO) as a fluor, with a concentration of $1.5\,\mathrm{g\,l^{-1}}$, to increase the light yield. Surrounding the IV is the outer nylon vessel (OV), which is itself inside of the Stainless Steel Sphere (SSS). The liquid between the IV, OV, and SSS, the *buffer*, is also a pseudocume solution, but without the addition of PPO and diluted with Dimethyl-phthalate in order to reduce its light yield - mainly to protect from gamma events originating in the PMT material and in the SSS as well as from fast neutrons. The OV served to protect from radon diffusion. This results in a total of about 1.3 kt of liquid and the SSS is additionally submerged in 2.4 kt of ultra-pure water to protect the detector from outside influence.

Borexino is equipped with 2212 inward facing 8" PMTs mounted on the SSS (ID). 208 outward-facing PMTs are additionally mounted on the SSS and the floor (OD) - using the water tank as a muon veto. There are about 6000 muons crossing Borexino every day - about half of which cross only the OD and the rest cross both the OD and the ID [189]. All-in-all, Borexino reaches a muon veto efficiency of 99.992% [190]. The muons reaching Borexino, which is situated under the comfortable 3700 m.w.e. at the Gran Sasso national laboratory[5], have an average energy of (283 ± 9) GeV [191] and their rate was measured to be $(3.432 \pm 0.003) \times 10^{-4}\,\mathrm{m^{-2}\,s^{-1}}$. Borexino originally had a light yield of 500 p.e./MeV, resulting in an energy resolution of about 5% at 1 MeV for events in the detector centre. Thanks to its rigorous calibration programme, Borexino also achieved a 9 cm position reconstruction uncertainty (on the x and y coordinates for similar events [192]). However, the number of active PMTs in Borexino has decreased from the about 2000 in 2012 to an average of 1238 in Phase-III, due to the ageing of the PMTs [68]. In Phase-III, this results in a slightly worse performance of 6% energy resolution and 11 cm position reconstruction uncertainty for events in the centre at 1 MeV.

As previously mentioned, a thermal insulation was installed on Borexino in 2015. The insulation consisted of about $900\,\mathrm{m^2}$ of mineral wool with two layers of 10 cm thickness each. An active temperature control system as well as many temperature sensors were also placed in strategic positions in the detector, the insulation, and the experimental hall [184].

[5] https://www.lngs.infn.it/en

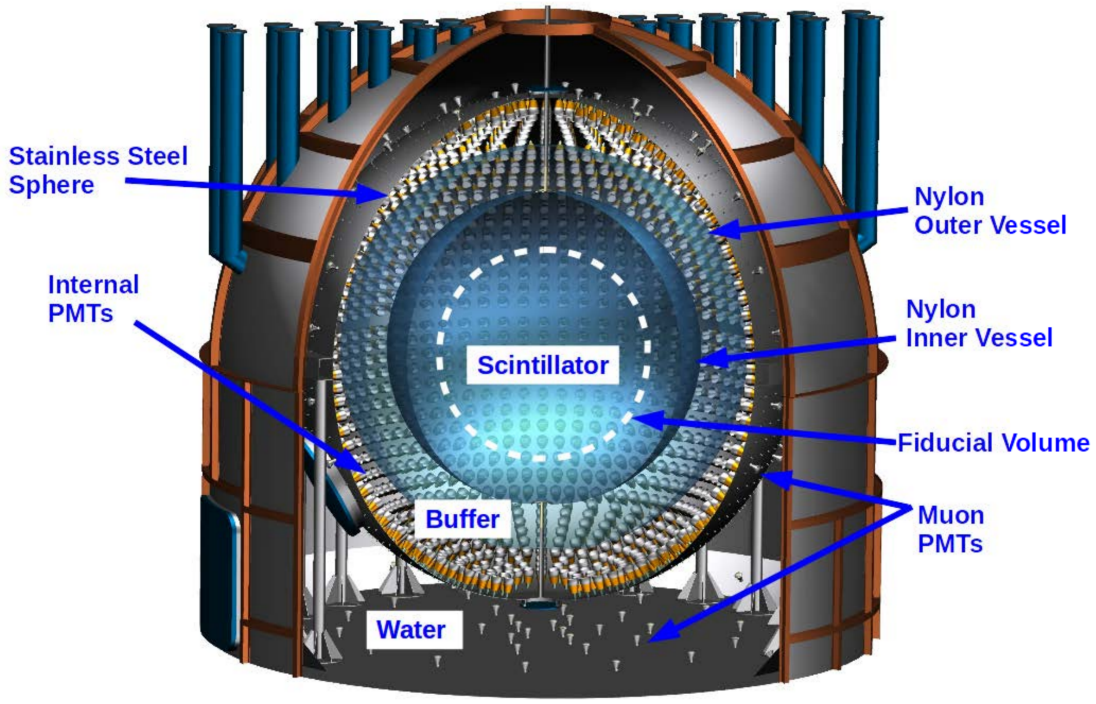


FIGURE 11: A sketch of the inside of the Borexino detector. In the centre, one can see the inner nylon vessel that holds the liquid scintillator and the outer nylon vessel that holds the buffer liquid (see text). The vessels are surrounded by more buffer liquid inside of the stainless steel sphere that also holds the PMTs. The entire sphere is itself submerged in ultra pure water. *Credit:* Borexino Collaboration [126]

At the time of writing, the draining of the scintillator of the Borexino experiment has started and the run-control electronics were turned off. This marks the end of a long and successful experimental life, but some final analyses remain to be published. This includes updates of the CNO analysis using Phase-III data beyond January 2020 and a separate CNO analysis without any ^{210}Bi constraint, using a combined Phase-II and Phase-III dataset, see section 3.7.4.

2.2 JUNO

The Jiangmen Underground Neutrino Observatory (JUNO) is a next-generation liquid scintillator experiment currently being built in the Guangdong province in China [156]. The main goal of JUNO, whose construction is scheduled for completion in 2023, is a 3σ determination of the neutrino mass hierarchy (MH) using reactor neutrinos within six years, as well as a precise measurement of θ_{12}, Δm^2_{21}, and Δm^2_{31}. As a multi-purpose experiment, owing to its unprecedented size and energy resolution, JUNO has many other goals reflecting the vast variety of sources whose neutrinos JUNO can detect. Section 2.2.1 goes over the many physics goals of JUNO and section 2.2.2 elaborates on the experiment's design. The expected rates from different sources, as well as the main backgrounds of the respective analyses, can be seen in table 4.

Neutrino source	Expected signal	Main background
Reactor	$60\,\mathrm{d}^{-1}$	Cosmogenics
Supernova (10 kpc)	5000 IBD + 2300 ES	Negligible
DSNB	$3\,\mathrm{yr}^{-1}$	Atmospheric neutrinos
Sun(^{8}B)	$16\,\mathrm{d}^{-1}$	Cosmogenics
Sun(*pep*,^{7}Be, CNO)	$568\,\mathrm{d}^{-1}$	Internal radioactivity
Atmosphere	$230\,\mathrm{yr}^{-1}$	Cosmic muons
Earth	$400\,\mathrm{yr}^{-1}$	Reactor neutrinos

TABLE 4: Different neutrino sources, their expected event rates in detected signals, and the main background contribution to the respective analyses. All values are approximate but give a good order of magnitude estimate, and are mostly based on Table 2 from Ref. [156].

The JUNO site is located about 215 km west-west-south of the Daya Bay nuclear power plant (NPP) and is placed at a distance of 53 km from both the Taishan and Yangjiang NPPs. It is hereby important that these two baselines are as equal as possible with a maximum of 0.5 km allowed in JUNO's design [92]. The Yangjiang NPP has six reactor cores, a total thermal power of 17.4 GW, and is operational since 2019. The Taishan NPP, where TAO is located (see section 2.2.2), was during the design phase of the JUNO experiment planned to have four reactor cores with a total thermal power of 18.4 GW, but only two of them were actually built. They are also operational since 2019. The entire experimental site was dug under a mountain, leaving the JUNO detector under 700 m of granite overburden or about 1900 meter-water-equivalent. The residual rate of muons in JUNO is expected to be $(10.8 \pm 1.1)\,\mathrm{m}^{-2}\,\mathrm{h}^{-1}$ and their average energy (209.2 ± 6.4) GeV [92].

2.2.1 Physics goals

Mass hierarchy As previously mentioned, JUNO's main goal is the determination of the neutrino mass hierarchy (see section 1.7), which is expected to achieve with a 3σ significance within six years of measurement. This will be achieved by measuring anti-electron-neutrinos from two nuclear power plants, the Taishan NPP and the Yangjiang NPP, each situated at a distance of 53 km from JUNO. This distance was of course especially chosen as it maximises the sensitivity to the neutrino mass hierarchy. The main strategy of the measurement, as can be seen in fig. 12, is to observe the effect of two oscillations simultaneously. This is possible because the 53 km baseline at which the L/E sits is in the first maximum of solar(Δm^2_{12})-driven oscillations. Because of the high frequency of the Δm^2_{13}-driven oscillations, a high energy resolution is required, as well as high statistics, which are the main drivers behind JUNO's enormous size and high light coverage.

JUNO detects reactor antineutrinos by measuring IBD events (see section 1.6.5) through a prompt-delayed coincidence signal. The two main sources of background events [92] are so-called *accidentals* from intrinsic radioactivity and cosmogenic background. *Accidentals* mimic IBD events when two events of energies similar to the prompt and delayed IBD events randomly happen to occur in the same time window that would be expected from IBDs. This only happens rarely, but the large amount of singles in JUNO means that it still affects the measurement. Cosmogenics are mostly decays of isotopes originating in the spallation of cosmic muon with (mostly) Carbon in the scintillator - whose mean lifetime exceeds reasonable muon veto cuts. The most relevant cosmogenic background comes from $\beta - n$ decays from ^{9}Li and ^{8}He. However, when muons pass near JUNO's water tank but don't (or only shortly) make direct contact with the water, fast neutrons are produced which can then scatter off a proton before being captured in the LS - thereby mimicking IBD signals. Further, α particles, which are mostly produced by ^{210}Po and the ^{238}U and ^{232}Th chains, can react with ^{13}C in the LS. $^{13}\mathrm{C}(\alpha, n)^{16}\mathrm{O}$ reactions can then also mimic IBD signals when excited ^{16}O emits a de-excitation gamma, or if the neutron is "fast" enough. Finally, geo-neutrinos, see section 1.5, are anti-electron neutrinos from radioactive decay processes inside of Earth's mantle and crust. They mainly affect the reactor spectrum in the lowest energy regions. An extensive sensitivity study was performed using simulations and all of the expected analysis backgrounds, with optimised cuts. These cuts reduce the rate of *accidentals* mainly by selecting events based on energy, time- and space-signatures.

Based on these studies, the collaboration formulated a contamination requirement of $1 \times 10^{-15}\,\mathrm{g\,g^{-1}}$ for ^{238}U and ^{232}Th in the liquid scintillation, which can be ensured by LS purification [193]. These contaminations will be monitored by OSIRIS during the

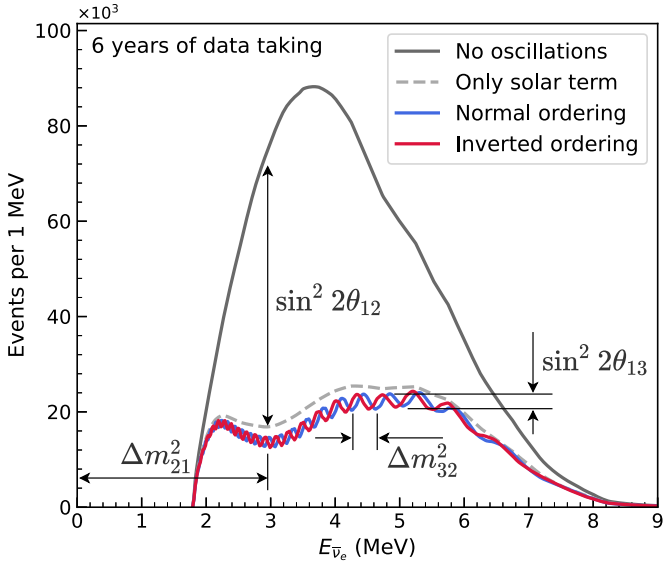


FIGURE 12: Expected reactor neutrino IBD energy spectrum for JUNO. The solid black line shows the unoscillated spectrum, and the dashed line shows the spectrum when only considering solar-driven oscillations. When adding θ_{13} oscillation, one either gets the blue or the red lines, depending on the MH. This difference is what JUNO aims to measure, for which it needs an unprecedented energy resolution. This figure also depicts how JUNO can be sensitive to four PMNS-matrix elements: namely $\Delta m^2_{21}, \Delta m^2_{32}$, $\sin^2\theta_{12}$, and $\sin^2\theta_{13}$.

filling of JUNO (see section 2.3). OSIRIS's sensitivity to ^{238}U and ^{232}Th is investigated in section 6. The two additional main requirements are an energy resolution of 3% at 1 MeV (stat+syst) and a non-linearity uncertainty of less than 1%. The latter is only achievable through a rigorous detector calibration, as outlined in [194]. Additionally, the Taishan Antineutrino Observatory (TAO, see section 2.2.2) will help by measuring an unoscillated spectrum from the Taishan NPP with an extremely low uncertainty, thereby quasi-nullifying the potential nefarious effects of unknown fine-structures in the reactor spectrum [195]. One should note that because JUNO's sensitivity to the MH comes from this oscillation interplay and not from matter effects, its data can be very fruitfully combined with that of other experiments (like PINGU [94]). JUNO's sensitivity will also slightly benefit from combining the reactor neutrino data with its own atmospheric neutrino measurements.

Precision measurements of PMNS-matrix elements As seen in fig. 12, JUNO will be sensitive to four PMNS matrix elements through its measured IBD energy spectrum. The study, which is under preparation for publication, also incorporates many details, such as the aforementioned background reduction strategies and considers e.g. an IBD cross-section that considers radiative corrections, shape information from TAO, a data-driven scintillator non-linearity model from the Daya Bay experiment, and more.

This results in an uncertainty of better than 0.6% on Δm^2_{31}, Δm^2_{21}, and $\sin^2\theta_{12}$ in six years, as well as sensitivity to $\sin^2\theta_{13}$ which is not better than existing measurements (PDG 2020) [156].

^{8}B solar neutrinos ^{8}B neutrinos are expected to be measured by ES in the energy range from about 2 MeV to 20 MeV, above the energies influenced by ^{11}C. The remaining ES backgrounds in that energy range are from reactor neutrinos whose rate declines rapidly with rising energy, cosmogenics which can mostly be tagged with muon veto cuts (there remain ^{10}C, ^{8}Li-^{6}He, ^{12}B, ^{11}Be in the energy range), externals which can be removed with fiducial cuts, and finally internal backgrounds (*i.e.* radioactive decays from contaminations in the LS). Assuming an internal contamination of $1 \times 10^{-17}\,\mathrm{g\,g^{-1}}$ for ^{232}Th and ^{238}U in the liquid scintillator, one can expect a total of 60.000 signal events with a 2:1 signal to noise ratio in ten years of data-taking.

Given the low energy threshold of the measurement and if $\Delta m^2_{21} = 7.5 \times 10^{-5}\,\mathrm{eV}^2$, as was measured by KamLAND [31], then a 3σ sensitivity on the absence of an "upturn" in the P_{ee} transition region can be achieved in the same time [196]. Additionally, the currently most precise measurement on the day-night asymmetry (1.1% by Super-K [47]) can be improved on. When using the spectral fit and the day-night asymmetry, one can simultaneous fit $\sin^2\theta_{12}$ and Δm^2_{21}, which results in an uncertainty on Δm^2_{21} which is on the same level as the current global fit results [197]. The inclusion of additional events from CC and NC interactions with ^{13}C nuclei, which could add up to 9.000 events in ten years, is under investigation.

Low energy solar neutrinos As explored in section 4, JUNO will also be able to measure solar neutrinos in the range from 0.45 MeV to 1.7 MeV using ES signals. This energy range includes recoil electrons from ^{7}Be, *pep*, and CNO neutrinos. The analysis of JUNO's sensitivity to those neutrinos is elaborated on in section 4 and was a substantial part of this thesis. The findings are currently being written in a to-be-published JUNO collaboration paper. *pp*-neutrinos are currently excluded from the analysis because of lacking simulation capacities in the low-energy region - particularly the trigger simulation. Additionally, pile-up is expected to play a large role in these energy regions - which requires a complex additional analysis. The great advantage of JUNO compared to other similar experiments [137, 198] is its unprecedented target mass and energy resolution. In about one year, JUNO is already expected to provide world-leading precision on the rates of ^{7}Be and *pep*-neutrinos, as well as measure CNO neutrinos (see section 4.5 for details).

Geo-neutrinos are electron antineutrinos from long-lived isotopes coming from inside our planet. Using a liquid scintillator detector like JUNO, one can detect IBD events from geo-neutrinos from the ^{232}Th and ^{238}U decay chains. JUNO has a unique position given its unprecedented size. In general 400 geo-neutrinos are expected to be measured per year [156]. Until now, the only two other experiments that have measured geo-neutrinos are KamLAND and Borexino, who respectively measured 169 neutrinos in six years and 53 in nine [109, 110]. Large geological studies are being undertaken near JUNO using seismographic stations to study gravitational anomalies which can give density information about the local crust [199, 200]. Because of the usual $1/r^2$ dependence on signal rates, this is the most important source of systematics for JUNO's capability to extract the mantle contribution of the geo-neutrino signal. This will allow JUNO to study the radiogenic contribution to the terrestrial heat production, with a roughly 5% expected precision on the geo-neutrino flux, a precision of 30% and 15% on the contributions from uranium and thorium, respectively, and to constrain the crust contribution with 18% precision [92, 111, 201].

Atmospheric neutrinos are created, along with many pions, muons, and kaons, when cosmic-rays (primarily high-energy protons) trigger showers in the atmosphere. Because of the angle-dependent matter effect, from the passage through Earth, atmospheric neutrinos can be used to probe the MSW effect and thus, give additional sensitivity to the main MH measurement. Additionally, because ν_μ neutrinos produce tracks with a large time profile compared to the purely leptonic ν_e neutrino showers, JUNO will be able to use time-of-flight-corrected timing information to discriminate between the two neutrino species, and reconstruct their energy spectra through an unfolding procedure. This will constitute the first sub-GeV measurement by a liquid scintillator detector and will give small ($\leq 2\sigma$) sensitivity to $\sin^2\theta_{23}$ and δ_{CP} [106].

Core-collapse supernovæ The JUNO detector is equipped with a special dedicated trigger scheme which can detect events with a 100 keV energy threshold. For a typical supernova occurring at a distance of 10 kpc, one expects about 10.000 events to be measured in different channels within 10 s. Because such supernovæ events are only expected to happen about once to thrice every century, many detectors on Earth will be awaiting such an event and JUNO will join the supernova early warning system (SNEWS 2.0) [103], therefore having to be ready to quickly recognise that a supernova event happened in order to send warnings to other observatories. JUNO will be able to measure SN neutrinos from all flavours, allowing us to study star parameters and physics, and is expected, through timing information alone, to provide an upper limit of (0.83 ± 0.24) eV (95%CL) on the absolute neutrino mass [104].

The diffuse supernova neutrino background is the integrated flux of all past supernovæ, see section 1.5. With simple distance calculations and extrapolations based on the rate of supernovæ in our galaxy, this flux can be estimated to be on the level of $10\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [106]. This corresponds to up to four measured events per year in JUNO by measuring IBD events with an energy well above that of reactor antineutrinos. The main background (in the analysis energy range above the reactor spectrum) stems from NC interaction of atmospheric neutrinos on ^{12}C atoms. The success of this measurement relies entirely on the liquid scintillator's pulse shape discrimination capabilities which can under reasonable assumptions considerably reduce the NC background. Should the mean energy of DSNB neutrinos be at least 15 MeV, JUNO is expected to reach a 3σ discovery in ten years [156].

Multi-messenger astronomy JUNO will also be equipped with a FPGA-based independent trigger system which will use the hardware aptitude of the chips to quickly sort background events out using only so-called Q-T pairs, *i.e.* quickly gathered time and charge information from waveforms with a neural network. The sorted events will be used to observe broadband transient events, which are expected to make JUNO a large player in the SNEWS 2.0 system for multi-messenger astronomy. The events will also be written to tape, allowing JUNO to look back for coincidences with signals from other telescopes.

Exotic searches and new physics JUNO is also expected to reach competitive levels in different searches for new physics. The prominent example is JUNO's search for proton decay, through a triple coincidence search where a proton decays into a Kaon and a neutrino, followed by a muon and an electron:

$$\begin{aligned} p &\to K^+ \bar{\nu} \\ K^+ &\to \mu^+ + \nu_\mu, \\ \mu^+ &\to e^+ + \nu_e + \bar{\nu}_\mu. \end{aligned} \tag{14}$$

The charged products are all potentially identifiable in liquid scintillator and thanks to its large target mass, JUNO is expected to reach a sensitivity of 8.34×10^{33} yr (90% CL) in ten years [156]. While going well beyond the scope of this thesis, JUNO will also be able to shed light on numerous topics beyond the standard model, including leading limits on Lorentz invariance violation, other nucleon decay modes, and spin-dependent dark matter-nucleon interaction cross sections in ranges below 20 GeV.

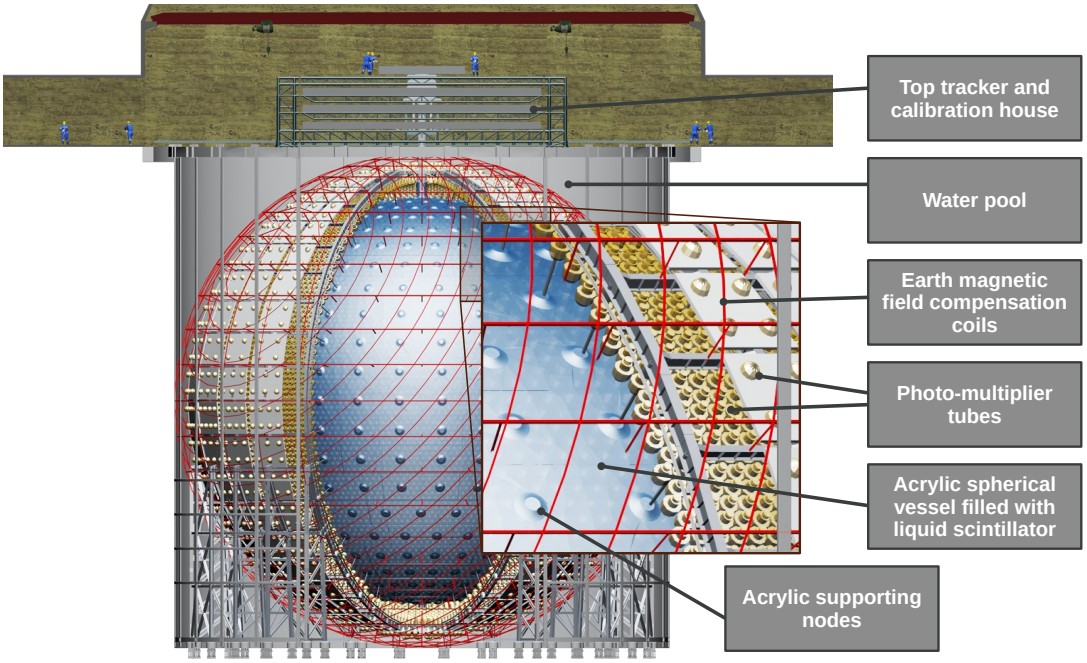


FIGURE 13: A 3-D sketch representation of the JUNO detector. *Credit:* JUNO collaboration

2.2.2 Detector design and status

The JUNO detector mainly consists of a 35.4 m diameter spherical acrylic vessel (AV) containing 20 kt of *linear alkylbenzene* (LAB)-based liquid scintillator (LS). LAB is a widely used solvent, in which 2,5-diphenyloxazole (PPO) is added to produce light yield. This happens because PPO molecules absorb excitation energy from the LAB and then produce photons per de-excitation. P-bis-(o-methylstyryl)-benzene (bis-MSB) is then also added to the scintillator in order to shift the photons to longer wavelengths, thereby preventing them from being re-absorbed by the LAB. This is crucial to allow photons to travel far enough in the large detector to reach the PMTs. The concentration of those added chemicals can however also be problematic as they have different absorption spectra. These concentrations were optimised in Daya Bay detectors in order to maximise the effective light yield in JUNO *i.e.* how much light is detected on the PMTs, and resulted in an optimal concentration of $2.5\,\mathrm{g\,l^{-1}}$ for PPO and $1\,\mathrm{mg\,l^{-1}}$ to $4\,\mathrm{mg\,l^{-1}}$ for bis-MSB [193]. The optical properties were also tested and it was shown that the then-employed purification substantially improved the transparency.

The AV itself is supported by a stainless steel lattice shell (SSLS) that is itself inside of a large cylindrical pool filled with ultra pure water. The AV is surrounded by 17612 20” PMTs (LPMT) and by 25600 3” PMTs (SPMT). The PMTs are attached to the SSLS and protected by magnetic coils which are meant to counteract Earth’s magnetic field. By putting the small PMTs between the large ones, JUNO is able to reach a light

coverage of about 78% (75.2% from LPMTs and 2.7% from SPMTs). Above the central parts of the water tank is the top tracker (TT), which is a muon tracker used to identify muons coming from the top and most useful to calibrate the muon track reconstruction in JUNO. This is an important factor because the muon veto cuts are currently projected to have an 83% signal efficiency in the MH analysis [156]. Additional 2.000 20" PMTs are placed inside of the water pool in order to serve as a muon veto, using Cherenkov light from the muons generated inside of the water pool. A sketch of the setup can be seen in fig. 13. Finally, at the centre of the TT lies a *calibration house* from which calibration sources will be automatically lowered inside of the AV. The calibration of JUNO will be performed with the combination of different systems. The Automated Calibration Unit [202] will be used to lower four different kinds of sources along the central axis of the detector. The sources include two neutron sources: Am-C and Am-Be, radioactive isotopes (namely ^{68}Ge, ^{137}Cs, ^{54}Mn, ^{60}Co, and ^{40}K), and a diffuser ball connected to a pulsed laser. The cable loop system will also place sources in the detector. It can reach many places along a vertical half-plane using the interplay of two cable spools. Then, a Guide Tube surrounding the outer side of the AV will also be used to transport sources in a longitudinal loop. Finally, a Remotely-Operated-Vehicle equipped with a camera and capable of manoeuvring in the scintillator, by using two pumps controlled through an "umbilical" cable, will also be able to carry a source and transport it anywhere in the detector [194].

A great amount of effort was put into the radiopurity of the JUNO detector. All used materials were carefully screened and selected, the production and cleaning procedures were optimised, as was the detector design - with constant control through the JUNO simulation software [203]. Additionally, the electronics of the JUNO experiment constitute a technological challenge. JUNO expects about 40 GB/s of triggered waveforms from all its PMTs. An online software (Online Event Classification) builds and processes the data in these waveforms to filter out relevant events and physical information - which reduces the data stream by about a factor 500 [156]. As previously mentioned, JUNO is also equipped with a dedicated supernova trigger system with a $\mathcal{O}(100\,\mathrm{keV})$ - JUNO is expected to be able to handle the about two million events (in a 10 s time window) expected from a supernova at 0.5 kpc. Additionally, JUNO is equipped with a parallel so-called multi-messenger-trigger with a $\mathcal{O}(10\,\mathrm{keV})$ threshold. It will constantly read data in a trigger-less mode and run fast algorithms to filter out dark noise - writing time/charge information to disk. The electronics and control soft- and hardware of JUNO communicate using a multi-layered EPICS[6] software. With an expected two petabytes of data per year, there is a large computing infrastructure behind the JUNO experiment. During normal operations, data on detector conditions, configuration files,

[6] https://epics-controls.org

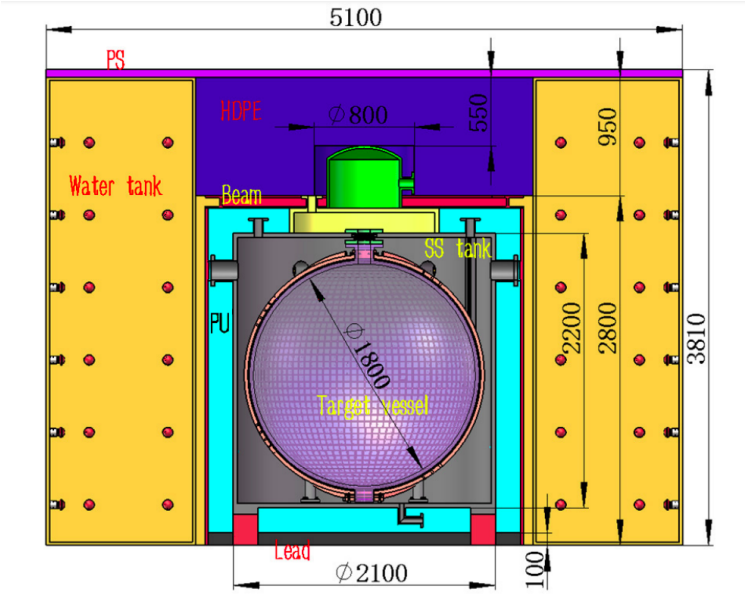


FIGURE 14: A sketch of the TAO detector. A central spherical acrylic vessel holds about 2.6 t of liquid scintillator cooled down to −50°. The outside of the acrylic is covered with SiPM detectors which usually have a high dark noise rate but not in such a cold environment. The detector is surrounded by an ultra-pure water tank to protect it from outside effects. Mostly because TAO is not placed under the protective cover of 700 m of granite, like JUNO, it is covered by high-density polyethylene (HDPE). *Source:* [132].

calibration data, and monitoring results will be accessible via a web service or CondDB [156]. The performance of the DAQ itself will also be monitored in real-time. The raw data will be transferred to the IHEP computing centre and then shared with other sites, including European sites, in a tiered computing system in which low-level and high-level analyses are performed in physically separate systems.

The JUNO experiment consists of two additional detectors: OSIRIS, which is a central focus in this thesis and is presented in more detail in section 2.3, and the Taishan Antineutrino Observatory (TAO) [132]. TAO is a cold gadolinium-doped liquid-scintillator experiment. It holds about 2.6 t of liquid scintillator in a spherical volume with a diameter of 1.8 m and operated at −50°. This temperature is advantageous because it allows the SiPMs, which are placed all around the acrylic vessel to reach a light coverage of 99%, to operate in a low-background mode. Finally, the acrylic vessel is surrounded by a stainless steel tank that is itself surrounded by a significant water pool to shield it from externals. A thick high-density polyethylene (HDPE) cover is situated above the AV to shield from ambient radioactivity and cosmogenic neutrons [132]. A sketch of the experiment's layout can be seen in fig. 14. With this high coverage, TAO will measure about 4500 p.e./MeV, resulting in an energy resolution of 1.7% at 1 MeV. The Gd-doped scintillator will allow a very efficient IBD detection (through better neutron tagging, see e.g. [38]). All in all, TAO, which is only 30 m away from a core of the Taishan nuclear

power plant, will provide a model-independent shape of the unoscillated reactor neutrino spectrum, which will be used to cancel systematics in the reactor shape uncertainty. It will also have its own physics programme, including an investigation of reactor neutrino anomalies [204, 205], investigating possible reactor fine structure [195], and looking for eV-scale sterile neutrinos.

2.3 OSIRIS

The Online Scintillator Internal Radioactivity Investigation System (OSIRIS) is a subdetector that is part of the liquid handling system of JUNO (see 2.2). Its main purpose is to monitor the radioactivity of the liquid scintillator during the filling of JUNO, particularly ^{238}U and ^{232}Th contaminations. It will also be able to search for air-leaks in JUNO's purification pipe system, measure the amounts of ^{14}C and ^{210}Po in the scintillator, and set constraints on other JUNO-relevant backgrounds (particularly for the solar analysis). Because of the large quantities (20 kt) of liquid, this process is expected to take about half a year and the properties of the liquid scintillator (which will be produced in batches at the JUNO site) are expected to vary over time. The JUNO collaboration has defined clear requirements concerning the radiopurity of the liquid scintillator: it can contain no more than $1 \times 10^{-15}\,\mathrm{g\,g^{-1}}$ of ^{238}U and ^{232}Th since any more would compromise JUNO's MH measurement [92]. Furthermore, the collaboration has deemed it desirable to reach levels of $1 \times 10^{-16}\,\mathrm{g\,g^{-1}}$, which would not only improve JUNO's MH measurement, but also make JUNO competitive in the landscape of solar neutrino measurements.

These extremely low radioactivity levels are not unprecedented in liquid scintillator neutrino experiments [46, 206], however the cumulative experience of past these experiments has indicated that even small problems can cause large contaminations in the scintillator. Air leaks in particular can cause relatively large quantities of *e.g.* radon and ^{85}Kr to enter the scintillator, and even past air exposure can cause large problems through emanation. The long time needed to fill the central detector translates to a high probability that at least one such error may occur. Since a single "bad batch" could cause serious harm to JUNO's physics goals, the fast identification of radioactive contaminants is crucial. However, it is quite difficult to actually measure such small contamination levels (at least over reasonable amounts of time), and standard methods like mass spectrometry or gamma spectroscopy are simply not applicable for this scenario. Thus the idea of the OSIRIS detector was born: a ton-scale liquid scintillator detector as a radioactivity monitor for JUNO - to detect even small contaminations during the filling process. As explained in Ref. [207], reaching even more extreme radiopurity levels, down

to $1 \times 10^{-17}\,\mathrm{g\,g^{-1}}$ in ^{238}U and ^{232}Th, is in general possible with OSIRIS, though not on the time-scales needed during the filling of JUNO.

2.3.1 Physics goals

As mentioned above, the main goal of the OSIRIS detector is to monitor the radiopurity of JUNO's liquid scintillator (LS) during the half-year long period of the filling of JUNO's central detector [92]. The detection of ^{238}U and ^{232}Th can be achieved within reasonable time scales by exploiting the fast coincidence of $^{212-214}$Bi-Po decays of the respective decay chains [207], as will be explored in section 6 of this thesis.

As the last element in JUNO's chain of liquid scintillator purification, OSIRIS is also the only guarantee that internal radioactive backgrounds don't exceed certain limits. Indeed, OSIRIS will not only be used to give stringent limits on ^{238}U and ^{232}Th, but also to precisely measure the rates of ^{14}C and ^{210}Po, and to give limits on other intrinsic backgrounds. All JUNO-relevant backgrounds that OSIRIS can set limits on, as well as the respective contamination requirements set by the JUNO collaboration, can be seen in Table 5. In case of an air-leak, JUNO will also be able to measure the rate of ^{85}Kr, which can in turn be used to set a limit on the rates of ^{39}Ar, see section 6.5.

Isotope	JUNO IBD (g/g)	JUNO solar (g/g)
^{238}U	10^{-15}	10^{-16}
^{232}Th	10^{-15}	10^{-16}
^{210}Po	N/A	$5 \cdot 10^{-24}$
^{40}K	10^{-16}	10^{-17}
^{14}C	10^{-17}	10^{-17}

TABLE 5: Liquid scintillator maximum contamination requirements for different isotopes set by the JUNO collaboration [92]. The two right columns refer to different scenarios that were tuned on IBD and solar neutrino measurements with JUNO, respectively.

2.3.2 Detector design

The OSIRIS detector is composed of concentric cylindrical shells holding different components. The entire detector is situated in JUNO's liquid scintillator hall, close to the central detector. The innermost part of OSIRIS is an acrylic vessel with 3 m diameter and height, containing about 21 m^3 (or 18 t) of liquid scintillator. The vessel, supported by eight acrylic plates, is mounted on a stainless steel frame that is itself submerged in ultra-pure water. The water, which acts as a buffer protecting the LS from external backgrounds from the photo-multiplier-tubes (PMTs) and the surrounding rock, is contained in a water tank with 9 m diameter and height. The stainless steel frame,

hereinafter referred to as the "PMT frame", of 7 m diameter and height, holds a total of 76 PMTs at a distance of 1.3 m from the AV. This distance was optimised to reduce the backgrounds induced by the radioactivity of the PMT glass while keeping the photo-coverage sufficiently high ($\approx 9\%$) for the Bi-Po analysis [208]. The setup can be seen in fig. 15. In addition to what can be seen in the figure, an array of thin PET sheets are spanned between the vertical rays of the PMT frame, also in between the iPMTs, and horizontally at the top and bottom of the tank. Black on the inside and white on the outside, this so-called *optical separation* prevents cross-talk between the *inner* and *outer* detector, which can be seen as logical volumes for the main analysis and muon veto, respectively. The reflections on the white outside planes of the optical separation were shown to enhance the performance of the muon-veto system [207].

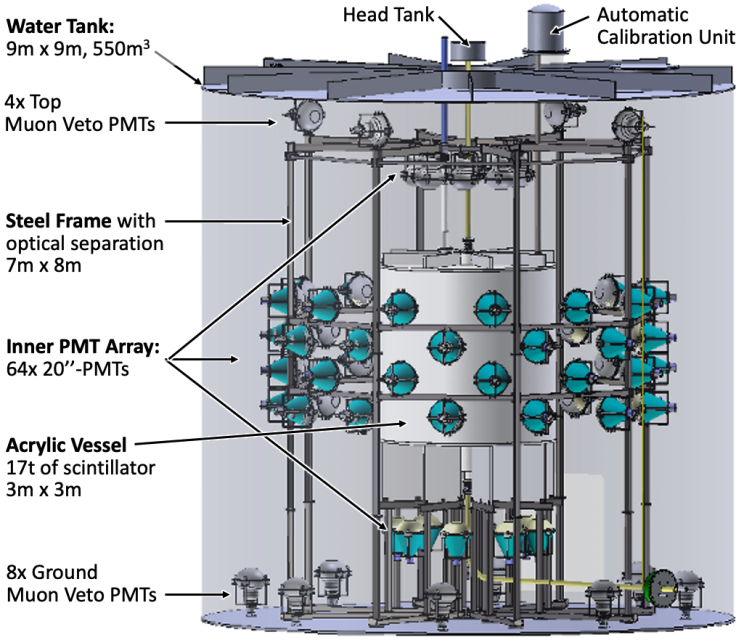


FIGURE 15: The OSIRIS detector. As indicated by the labels, here one can see the concentric cylindrical layout of OSIRIS, with the acrylic vessel (AV) holding the LS in the centre of the detector, itself submerged by ultra-pure water. The PMT frame (see text) is also visible, as are the 64 iPMTs surrounding the AV and the twelve veto iPMTs. One can also see the automated calibration unit resting on the top of the water tank, as well as different pipes and flanges going through the detector, including the calibration pipe that is used to transport sources directly inside of the liquid scintillator volume. *Source:* [207].

The so-called *intelligent PMTs* (iPMTs) used in the OSIRIS experiment are 20" Hamamatsu PMTs that were outfitted with a special electronic base each, see fig. 16. In this design, each PMT holds its own high-voltage source, digitiser, and control, all mounted inside of a stainless steel shell to avoid contact between the electronic components and the ultra-pure water. This setup also naturally increases the quality of the signal by significantly reducing the length of the cables between the raw PMT output and the

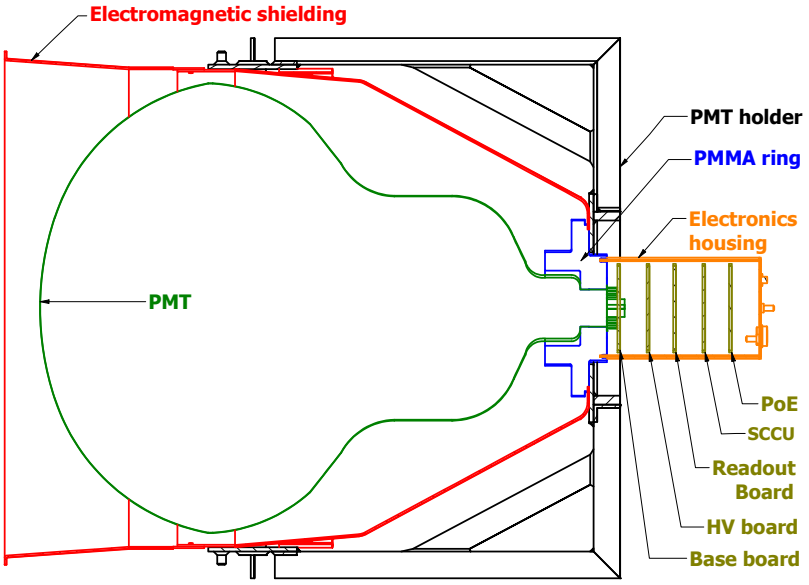


FIGURE 16: Sketch of an iPMT. The *intelligent* part of the PMT is in the base, which is connected to a single PoE-cable which simultaneously powers the iPMT and transfers the digitised PMT output to the OSIRIS DAQ, see text. The iPMTs were entirely designed (and produced) by a group at the Physikalisches Institut 3.b. at the RWTH Aachen. *Source*: [207].

digitisation. The digitised output is then autonomously transferred from the thus *self-triggered* iPMTs via Power-over-Ethernet (PoE) cables - which simultaneously serve as power supply. The digitised waveforms from the iPMTs sampled from the ADC with VULCAN chips, which are produced in the ZEA-2 institute of the Forschungszentrum Jülich GmbH. They can sample the ADC with a rate of 500 MS/s (resulting in 2 ns time-bins) in three different gain channels. Additionally, the iPMTs are equipped with a *Field Programmable Gate Array* on which waveform reconstruction can already occur. The data from all PMTs is collected and sent to the so-called *event builder* - an online software that filters out dark noise by trying to keep as much physics-relevant data as possible. This reduces a data stream of about 230 MB/s to an expected roughly 170 kB/s. This output data is stored on-site and sent over network to the IHEP data centre. It is also processed by an online software framework called the *RootSorter* [209] which was modified for use in OSIRIS. This framework is used to process and display data in real-time as well as possibly create alerts, which is particularly important for OSIRIS as *e.g.* possible air-leaks or ^{238}U/^{232}Th contaminations must be identified as quickly as possible.

Resting on the PMT frame, the iPMTs are further equipped with cone-shaped magnetic shields that protrude out the front side of the iPMTs. As can be seen in fig. 15, 64 iPMTs are placed around the acrylic vessel and facing inward, and twelve additional

iPMTs are placed at the top and bottom of the detector. The latter serves as a muon veto, using the water buffer as a Cherenkov detector. In order to preserve a laminar flow during the *continuous* operation of the detector, a diffuser is placed at the top centre of the acrylic vessel. This diffuser is designed to turn the vertical motion of the liquid flux horizontally while preserving a laminar flow. This ensures minimal mixing of new and existing LS. For the same reasons, the temperature of the new LS entering from the top of the AV is set to be 4° higher than the LS at the bottom, as a temperature gradient stabilises vertical flow. The setup was tested on a 1:10 scale prototype [207].

Similarly to JUNO, OSIRIS will be equipped with an ACU to perform source calibrations. The design of and further work behind OSIRIS's source calibration system is described in full in section 5 and is a major topic of this thesis. Another calibration system, the so-called *light-injection-system*, also exists. It consists of a pulsed pico-second laser and a network of optical fibres that deliver said pulses to different diffuser balls installed on the PMT frame. This system was mainly developed by Tobias Sterr[7], can perform very precise timing and charge calibration of the iPMTs, and is briefly described in section 5.3.

OSIRIS Simulation All of the physics and design preparations for OSIRIS, including the work on the calibration, was performed alongside the development of and with the constant input of a Geant4-based simulation software. OSIRIS's simulation is a Geant4.10.02.p03-based C++11 Monte-Carlo framework, which was built by Christoph Genster[8], Paul Hackspacher[9], Sebastian Lorenz[8], and later myself in the context of calibration-related work for this thesis. It describes OSIRIS's geometry in detail and takes into account a realistic behaviour for light particles in the scintillator, as was done in the JUNO offline software [92, 208]. The OSIRIS simulation does not take into account simulation of OSIRIS's electronics, but can output the number and timestamp of each photon hit on a PMT for each event. It has different modes, including a *Calib* mode in which radioactive sources are placed in the detector, and *LEDCalib* in which a LED is placed. The latter also implements the angular behaviour of the light coming from the LED capsule, as was measured in a setup in Tübingen (see appendix D).

[7] Eberhard Karls Universität Tübingen
[8] IKP-2 Forschungszentrum Jülich GmbH
[9] Johannes Gutenberg-Universität Mainz

3 Solar neutrino measurements with Borexino

The measurement of neutrinos produced in the CNO nuclear chain in the Sun, as detailed in section 1.3.3, is from an experimental perspective very challenging. Indeed, the interaction rate of these neutrino events is of the order of just a few counts per day per 100 t of target mass. This is small even compared to the other solar neutrinos and, additionally, the expected energy spectrum of CNO neutrinos has no prominent features that could make detection easier. The measurement of CNO neutrinos is however also extremely interesting for several other reasons, see section 1.3.3. The most important are that 1) the CNO process was never directly observed and 2) because a measurement of the rate of CNO neutrinos could lead to the solving of the solar metallicity problem, a long standing problem in solar physics. While Borexino's sensitivity to measuring the CNO solar neutrino rate is not sufficient to discriminate between the LZ and HZ models of solar metallicity [60], a first direct experimental discovery of CNO neutrinos in the Sun sets a big milestone in a better understanding of our Sun, and of stars in general.

In the following, sections 3.1–3.4 elucidate how solar neutrinos are usually detected in Borexino and what special considerations were taken in the case of CNO neutrinos, sections 3.5 and 3.6 how the ^{210}Bi constraint for the fit was obtained from ^{210}Po data, and finally section 3.7 gives the results.

3.1 Solar neutrino analyses with Borexino

Borexino can detect solar neutrinos through their elastic scattering off electrons in its liquid scintillator. Because the so-produced flashes of light are in general indistinguishable from backgrounds (intrinsic radioactivity, cosmogenic isotopes, muons, intrinsic radioactivity) on an event-by-event basis, different shape analyses [46, 183] were used to statistically extract relevant physical quantities from background.

Unlike simpler ,,counting experiment" approaches, spectral fits do not only employ the total number of events in given bins, but also incorporate a-priori information about the spectral shapes of different contributions, and/or auxiliary measurements. In this case, the main output of a fit is an array ($\vec{\theta} = \theta_0, .., \theta_n$) of the total rates of the different contributions (neutrino species, backgrounds, detector parameters in some cases) and an array ($\vec{\alpha} = \alpha_0, .., \alpha_m$) of nuisance parameters usually related to the detector response. While in general many different methods can be used to improve the performance of such fits (e.g. Borexino simultaneously fits energy spectra, pulse-shape data, and radial distributions [46] to greatly improve fit stability and reduce systematics), the success of said fits mainly relies on two factors. Firstly, some shapes have certain distinctive

features that make them clearly identifiable / stand out amongst other shapes (e.g. the ^{7}Be shoulder). These features end up acting somewhat like standard candles, automatically constraining that parameter (θ_i) and the sum of all other contributions in a given spectral range. Secondly, external constraints can be set on *e.g.* the total rates of some of the contributions or on the ratio of two contributions, to break degeneracies where they exist.

For example, in Borexino, the shapes of pile-up events and *pp*-neutrinos are very similar in the spectral range where the noise-to-signal ratio is minimal for *pp*-neutrinos, from about 230 keV to 280 keV). Because these shapes do not significantly influence other energy ranges, where constraints could be found, this makes it in principle impossible for a spectral fit to disentangle both components. This would therefore result in very high anti-correlations and general unstable behaviour. However, independent measurements (from second-cluster data) of the ^{14}C rate[46] make it possible to externally constrain the rate of pile-up (see appendix A). This breaks the existing degeneracy between *pp*-neutrinos and pile-up, and allows the fit to provide a reliable measurement of the rate of *pp*-neutrinos.

Two approaches were used by the Borexino collaboration in past publications. The first is called the *analytical fit*. The idea is to model all particle interactions mathematically in order to reconstruct the spectra of different events in the Borexino detector, as a function of the particle interaction rates of the different signal and background species, but also of different detector properties such as the quenching factor and the light yield. While the high number of free parameters introduces obvious difficulties, this approach has the added benefit of providing information about the stability of the detector and general cross-checks of how well the detector is understood. In this approach, the theoretical solar neutrino energy spectra are firstly convoluted with the electron recoil cross-section, then the spectrum is converted to the N_p variable (number of fired PMTs per event). The second approach is called the *Monte-Carlo fit*. In this approach each signal and background species is simulated using Borexino's MC, see section 3.2, and the corresponding N_{hits} energy spectra PDFs are saved. These spectra are then fitted to the given data sample by simply scaling the PDFs to the data, thereby directly obtaining rates for each fit species.

3.2 Monte-Carlo software and data

The simulation software of the Borexino experiment is a Geant4-based C++ software [210, 211]. It simulates events using a software chain designed to output results in the same format as the real data taken by Borexino, so that the exact same analysis

software may be used on both data and MC events. The simulation is closely tuned to the calibration runs performed with Borexino [192]. This was done by comparing the energy depositions of several calibration sources at different positions, giving reference points for different particle types, energies, and positions. It allows to understand the uniformity of the energy response, the detector stability over time, and the effects of different reconstructions.

The process of creating the MC data is also closely tuned to the data-taking procedures: for each data run, the simulation is provided with all relevant time-dependent physical inputs *e.g.* run duration, effective quantum efficiency [189], vessel shape, and live channels to simulate events with a correct time-dependence on the detector properties. The number of simulated events is also proportional to the duration of each data-taking run. While this approach is less sensitive to the stability of the detector properties over time, it has the obvious benefit of requiring less nuisance parameters, thereby permitting a stabler, faster minimising procedure in the fit.

In some more detail, physical events and processes are first simulated, and the produced light is propagated through the liquid scintillator and detector geometry using Geant4. This continues for each particle and created sub-particle until a certain lower energy threshold is attained or until each particle has been absorbed. For the PMTs that were hit by at least one photon in this first simulation step, a subsequent electronic detection is then simulated through a software chain. Finally, the results undergo the same reconstruction steps and selection cuts as real data, and the surviving events form the particle spectra used in the fit. The energy estimator used in this approach is N_{hits}^{norm} *i.e.* the total number of triggered hits on all PMTs for any given readout window, normalised to 2000 PMTs.

3.3 Measuring *pp*, *pep*, and ^{7}Be solar neutrinos

As outlined in section 2.1.1, Borexino measured not only the rate of ^{7}Be neutrinos (its original goal), but also of ^{8}B, *pp*, and *pep* neutrinos [46, 212]. This section broadly describes how they were measured as the analysis methods tie-in to the CNO discovery.

3.3.1 High Energy Region (HER) - ^{8}B neutrinos

The HER analysis occurs from 3.2 MeV to 16 MeV. Unlike its low-energy counterpart (see below), this energy region is not affected by long-lived radioactive backgrounds (with the exception of ^{208}Tl in the HER-I, see below). The dataset is further divided into HER-I and HER-II, cutting around 5.7 MeV because there is no background from

natural radioactivity above this energy. For these reasons, a large dataset can be used, with data from 2008 to 2016. While the details are beyond the scope of this thesis - it is worth noting that the FV cuts in the HER are in general also more relaxed than those used in the LER, and that there are differences in the cuts even between HER-I and HER-II events, *e.g.* the fiducial volume cut which in the latter case includes the complete IV volume.

In order to extract the rate of ^{8}B neutrinos from the backgrounds, the radial distribution of the events in the HER-I and HER-II were fitted separately and the results combined. Since spectral distortions due to neutrino flavour conversion have no impact on the shape of the radial distributions, this fit can be performed independently of any assumptions on the energy shape of ^{8}B neutrinos or neutrino oscillations.

3.3.2 Low Energy Region (LER) - *pp*-chain neutrinos

The LER spans events from 0.19 MeV to 2.93 MeV and from 2011 to 2016, the so-called *Phase-II* dataset. The data then undergoes a series of cuts mainly directed at removing muon events, external backgrounds, but also regarding data-quality. Because of the irregular shape of the nylon vessel, the fiducial cut is defined as:

$$r < 2.8\,\mathrm{m}\ \&\ 1.8\,\mathrm{m} < z < 2.2\,\mathrm{m}, \tag{15}$$

where r, z are the reconstructed events radius and z-coordinate, respectively. The asymmetry in the z-cut is mainly related to radioactive background from the endcap region of the nylon vessel. Another notable cut is the so-called TFC-cut which relies on the (time and space) three-fold-coincidence [213] between a muon event, a positron event from cosmogenic ^{11}C and neutron capture. This cut is used to separate the data in two distinct datasets: a TFC-tagged dataset enriched in ^{11}C events, and a TFC-subtracted dataset with very little ^{11}C remaining.

The data are then fitted with a multivariate binned Poisson likelihood fit in two approaches. The first is called the *Analytical fit*. The second approach, the *Monte-Carlo fit* (also the one used in section 3.7.2), relies entirely on the MC simulation of the Borexino detector [211]. The obvious advantage of this approach is the simplicity of the fit and the lower number of nuisance parameters. Both approaches are described in more detail in section 3.2. The multivariate fit performed in both approaches simultaneously fits TFC-tagged and TFC-subtracted datasets, the radial distribution of events, as well as a pulse shape parameter. The separation of the TFC-subtracted and TFC-tagged

datasets makes a sensitivity to *pep* neutrinos possible, the radial distribution can disentangle internal from external events, and the pulse shape distribution, which is only applied to the TFC-subtracted spectrum, serves to further disentangle ^{11}C events.

3.3.3 Results

The resulting solar neutrino rates as measured by Borexino can be seen in Table 6. These results represent the first simultaneous spectroscopy of all solar neutrino species from the *pp*-chain. The findings constitute an important confirmation of the SSM and were used to confirm that the Sun has been in thermodynamic equilibrium for at least 1×10^5 yr, to provide a (weak) hint towards the HZ model of solar metallicity, and to study relative *pp*-chain termination rates. They can also be used to test the MSW-LMA paradigm as the different neutrino species have different energies, therefore making a P_{ee} measurement possible as a function of energy. This measurement can be performed by Borexino over the entire vacuum transition region when including ^{8}B measurements. In general, the findings are somewhat model-dependent as the conversion from event rates in the detector to neutrino fluxes from the Sun depends on oscillation parameters, the number of target electrons in Borexino, the MSW-LMA effect and, in the case of *pep* neutrinos, on the metallicity, since the rate of CNO neutrinos in the fit has to be constrained to the SSM prediction.

Neutrino species	Flux ($\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$)
pp	$(6.1 \pm 0.5^{+0.3}_{-0.5}) \cdot 10^{10}$
^{7}Be	$(4.99 \pm 0.11^{+0.06}_{-0.08}) \cdot 10^{9}$
pep (HZ)	$(1.27 \pm 0.19^{+0.08}_{-0.12}) \cdot 10^{8}$
pep (HZ)	$(1.39 \pm 0.19^{+0.08}_{-0.13}) \cdot 10^{8}$
^{8}B	$(5.68^{+0.39+0.03}_{-0.41-0.03}) \cdot 10^{6}$

TABLE 6: Results from the Borexino solar analysis [46] converted to fluxes (with model-dependent oscillation assumptions and for the estimated number of target electrons). All fluxes are integral values without any threshold. The *pep* neutrino result depends on whether we assume HZ or LZ SSM predictions to constrain the CNO neutrino rate.

3.4 CNO neutrinos in the Borexino experiment

The spectral region most sensitive to CNO neutrinos extends from 780 keV to 885 keV [68] (see Figure 17). In the same energy range, the radioactive decays of intrinsic ^{210}Bi in the liquid scintillator, as well as *pep* neutrinos, constitute the main sources of background for this analysis. As can be seen in Figure 17, the spectral shapes of both of those components are very similar to the expected shape from CNO neutrinos, and they have comparable expected rates. Therefore a similar problem to the pile-up and *pp*-neutrino

example occurs in the pursuit of the discovery of CNO neutrinos: while it is possible to provide a broad upper limit on the rate of CNO neutrinos by looking at the sum of the three contributions, a real measurement can only be made possible through external constraints on the rates of both pep-neutrino events and ^{210}Bi decays.

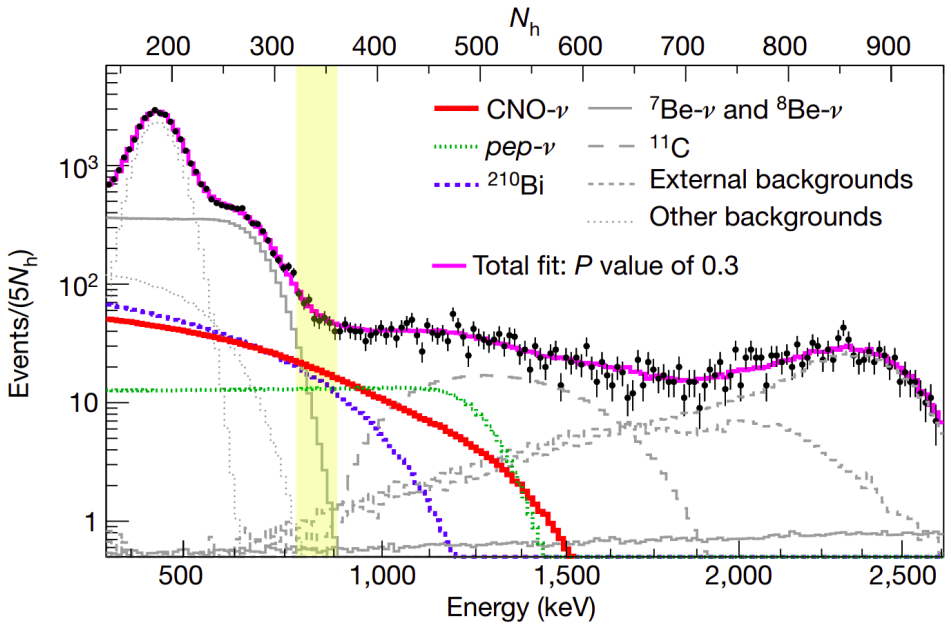


FIGURE 17: Borexino's measured energy spectrum and results of a spectral fit (explained in more detail in section 3.7.2). The solid red, dotted green, and dashed blue lines depict the spectra of CNO neutrinos, *pep* neutrinos, and ^{210}Bi decays, respectively. The energy region most sensitive to CNO is shown by the yellow band. The x axis shows the energy in keV (bottom) and in the N_h estimator (top).

For pep-neutrinos, finding an external constraint is relatively straightforward. First, one can look at the reactions generating *pp*-neutrinos and *pep* neutrinos:

$$(pp)\ p + p \rightarrow d + e^+ + \nu_e,$$
$$(pep)\ p + e^- + p \rightarrow^2 \mathrm{H} + \nu_e.$$

Because the same nuclear matrix element determines the probability of electron capture and β decay, the ratio of the fluxes $\mathrm{R}_p = \Phi(pp)/\Phi(pep)$ does not depend on nuclear physics, but only on the available phase space volumes and the electron density (of the solar interior). However, there are some kinematic differences between the *pep* electron capture and *pp* β decay which lead to small differences due to radiative corrections [214]. Treating these differences as systematics, this ratio can still be determined from theory with an uncertainty of about one percent. Combined with Borexino's previous measurement of the *pp*-neutrino rate [183], one can obtain a constraint on the rate of pep-neutrinos in Borexino with an uncertainty of about ten percent (one should note that *rate* refers to the rate of neutrino interactions in the detector, and *flux* to the incoming

flux from the Sun). Additionally, one can use photon luminosity measurements as the so-called ,,solar luminosity constraint" [215, 216] to improve this constraint by an order of magnitude (less uncertainty). To calculate the ratio (R_p), as described in [217], the approach from [218, 219] was used where the electron neutrino survival probability $P_{ee}^{2\nu}$ was firstly calculated for the very simple two-neutrino mixing case, and then extended to a three-neutrino framework using:

$$P_{ee}^{3\nu}(E) = \sin^4\theta_{13} + \cos^4\theta_{13} \cdot P_{ee}^{2\nu}(E), \tag{16}$$

where E is the neutrino energy. The total survival probability is then:

$$P_{ee} = \int_{E_{min}}^{E_{max}} \lambda(E) P_{ee}^{3\nu}(E) dE, \tag{17}$$

where $\lambda(E)$ is the corresponding solar neutrino energy spectrum. This survival probability can then be used to convert fluxes to rates in the Borexino detector:

$$R_\nu = n_e \Phi_\nu \int_{E_{min}}^{E_{max}} \lambda(E) dE \cdot \int_{T_{min}}^{T_{max}} \left[\frac{d\sigma_e^\nu(E,T)}{dT} P_{ee} + \frac{d\sigma_{\mu,\tau}^\nu(E,T)}{dT}(1 - P_{ee}) \right] dT. \tag{18}$$

Here, n_e is Borexino's electron density [181], T is the recoil electron energy, σ_X^ν is the interaction cross-section of X-flavoured neutrinos, and Φ_ν is the initial expected neutrino flux. Using Φ_ν from [220], neutrino mixing parameters from a global fit [221], cross sections as listed in Tab. 1 from [181], and the solar spectra ($\lambda(E)$) from John Bahcall's homepage[10], one obtains $R_p = 47.76 \pm 0.84$ for the HZ SSM. This, combined with the previous considerations, corresponds to an uncertainty of about 1.4% on the rate of pep-neutrinos, a very strong constraint to use for the CNO analysis.

Thus, in the following, the measurement of CNO neutrinos through a multivariate spectral fit is entirely dependent on the possibility of finding an external constraint on the rate of ^{210}Bi decays in the liquid scintillator. Compared to the determination of the *pep*-neutrino constraint, a more direct experimental approach is needed, which will be described in depth in sections 3.5 and 3.6.

3.5 The Low-Polonium Field (LPoF)

^{210}Bi has a relatively low half-life (five days). Its presence in Borexino's liquid scintillator can be explained by the presence of ^{210}Pb, which has a significantly higher half-life (22.3 years). A stable and homogeneously distributed amount of ^{210}Pb in the liquid

[10] sns.ias.edu/~jnb

scintillator would therefore result in a stable rate of ^{210}Bi decays through secular equilibrium. Because Borexino can in general not distinguish between neutrino signals or β decays on an event-by-event basis, it is also not possible to directly measure the rate of ^{210}Bi. However, as suggested in [222], demonstrating the existence of a region with a stable β-like signal rate inside of a liquid scintillator detector can lead to a measurement of the rate of ^{210}Bi through a measurement of ^{210}Po by virtue of the following decay chain:

$$^{210}\text{Pb} \xrightarrow[32.2y]{\beta} {}^{210}\text{Bi} \xrightarrow[7.23d]{\beta} {}^{210}\text{Po} \xrightarrow[199.6d]{\alpha} {}^{206}\text{Pb(stable)}.$$

This measurement is possible on an event-by-event basis because ^{210}Po undergoes an α decay that can be detected very efficiently using pulse-shape discrimination methods (in this case, a Multi-Layer-Perceptron or MLP).

In Borexino, the possible presence of such a region has been studied extensively. It was discovered that there are three sources of ^{210}Po in the liquid scintillator [60]. Firstly, **unsupported** ^{210}Po which was brought into the scintillator during different detector operations in 2007 and 2011. This component does not need to be discussed in this analysis (using data from after July 2016) because it was observed that it has since all but decayed. Secondly, a **scintillator** $^{210}\text{Po}^S$ component that is present in the liquid scintillator and in secular equilibrium with ^{210}Bi. Lastly, a **vessel** $^{210}\text{Po}^V$ which is understood to originate from ^{210}Pb on the inner surface of the vessel. The ^{210}Pb itself cannot be detaching from the surface, because measurements show a very stable rate of β decays in Borexino, but when the ^{210}Pb does decay its daughter-isotope ^{210}Po can detach and move into the scintillator before decaying [68]. Through diffusion alone, studies [223] show that the $^{210}\text{Po}^V$ can only travel about 20 cm before decaying. It can therefore not reach the FV which is one meter away from the vessel. However, temperature variations may induce convective motions inside of the liquid scintillator that may carry this $^{210}\text{Po}^V$ anywhere inside of the FV. As can be seen in Figure 18, there were very large movements of $^{210}\text{Po}^V$ inside of the detector before 2016.

Looking at the figure, one can define four distinct regions for the following analysis. The first region is from the start of the dataset (in 2016) to about mid-2017, where increased Polonium activity can still be seen near the centre of the detector (see the vertical line 3). The second is in 2019 (between the vertical lines 4 and 5), the so-called "disruption period" where a human error caused a sudden jump in the rates of polonium and a movement of the LPoF. The other two periods are between the two aforementioned ones and after the latter, where the detector and LPoF are stable.

These large movements stopped when, at the end of 2015, an extensive insulation campaign was performed [184]. It comprised not only thermally insulating the detector but

also installing a temperature control system. This greatly reduced the temperature variations in the detector and allowed the establishment of a positive vertical temperature gradient (of about $0.5\,^{\circ}\mathrm{C\,m^{-1}}$). With this new-found stability in the detector, one can expect a stable region, free of ^{210}Po from the vessel, to establish itself near the centre of the detector (slightly above the equator), as was shown by fluid simulations [223].

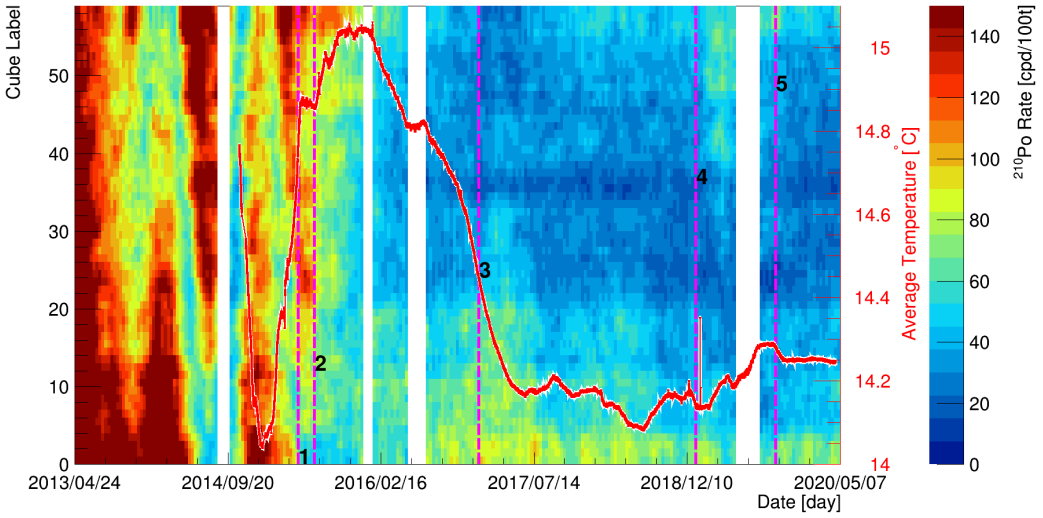


FIGURE 18: The time evolution of ^{210}Po decays in different z-slices of Borexino. The colour bar indicates the measured rate of ^{210}Po decays in a given bin, and the overlaid red line depicts the average temperature measured by the different sensors around the stainless steel sphere. The white bands indicate when data-taking was not working and the numbered dashed lines show the moments of different operations on the temperature control system. 1. Beginning of the insulation installation 2. Turning off water loop in water tank 3. Completion of insulation installation 4. Start of air temperature control system (ACTS) 5. Adjustment of ACTS set-points 6. Start of ACTS in experimental hall C. As one can see, the rates of polonium vary greatly before the installation of the temperature control system, and zones of low-polonium activity quickly form once it is active. The temperature also shows order-of-magnitudes less variation.

We therefore expect, over time, that a region will form which will contain only the $^{210}\mathrm{Po}^{S}$ component. A measurement of the rate of α decays inside of this so-called ,,low-polonium field (LPoF)" can be seen as providing a direct external constraint on the rate of ^{210}Bi decays, thereby making a discovery of CNO neutrinos in the Sun possible (see Sec. 3.4). The next sections explain how to perform this measurement.

3.5.1 ^{210}Po Data

The ^{210}Po data for this analysis were selected from July 2016 to February 2020 with relatively simple cuts on MLP and geo-normalized charge (total PMT charge normalised to 2000 PMTs with an additional weight proportional to the solid angle between the event cluster and the corresponding PMT) that ensure that α-like events are selected

in the well-known sharp energy range of ^{210}Po decays. Using the Monte-Carlo, the MLP efficiency for ^{210}Po was determined to be 96.9%. Additionally, the amount of ^{210}Bi events (β events with a similar energy) that ,,leak" into the selected energy region was calculated to be 2.6 cpd/100t. These quantities will hereinafter be referred to as *efficiency* and *leak*. The performance of the MLP pulse-shape discriminator can be seen in Figure 19 where it was verified with a ^{214}Bi-Po dataset [110].

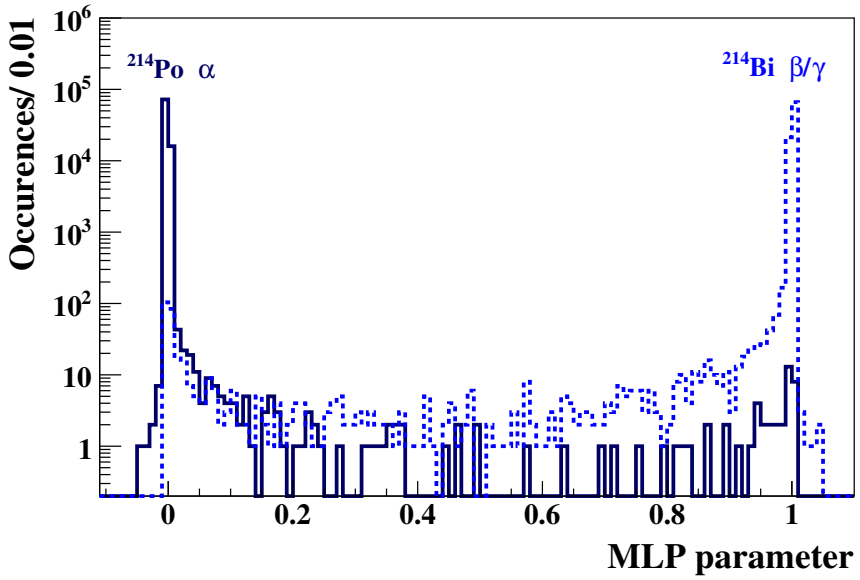


FIGURE 19: The distribution of the MLP pulse-shape discrimination variable on ^{214}Po and ^{214}Bi data [110]. As one can see, there is a very strong discriminatory power between α and β like decays in Borexino.

3.6 Extracting the ^{210}Bi constraint

Several approaches were explored to extract the ^{210}Bi constraint from the low-polonium field data. The following sections describe and compare them.

3.6.1 Paraboidal basis

During the formation of a low-polonium-field, given an ideal Borexino-like spherical detector in perfectly stable thermal conditions, one would expect the rate of ^{210}Po to exponentially decrease over the radial direction from the vessel to the centre of the detector. Let (0, 0, 0) be the centre of the detector in a Cartesian coordinate system. When projecting the ^{210}Po rate over a line going through the centre, one would therefore expect to see a parabola centered at zero. If one could fit this parabola, one could reliably extract the minimal rate R_{min} of the ^{210}Po decays in the detector. Since R_{min} is the

sum of two contributions (see Sec. 3.5), R_{min} can be interpreted as an upper limit on the rate of **scintillator** ^{210}Bi in the detector. This can in turn be used as a lower limit on the rate of CNO neutrinos (see Sec. 3.4).

After an extended period of time, in (perfectly) stable thermal conditions, the residual **vessel** polonium in the centre-most regions of the detector would slowly disappear according to the law of radioactive decay. Thus, a steadily-expanding region would form, in which ^{210}Po decays happen with a constant rate, therefore deviating from the previously expected parabola. In this scenario, a measurement of this plateau rate would directly lead to a measurement of the rate of CNO interactions in Borexino.

After these considerations, a straightforward approach can be applied: fitting a three-dimensional hyper-paraboloid through ^{210}Po data. In the following, an approach is described which only uses these simple considerations to fit the data. This approach is used as a basis and reference for the other approaches described further below. The following equation was used:

$$\frac{d^2R(^{210}Po)}{d\vec{x}} \propto (R_{min} * eff * geo + leak), \tag{19}$$
$$\text{with } geo = 1 + \left(\frac{x - x_0}{a}\right)^2 + \left(\frac{y - y_0}{b}\right)^2 + \left(\frac{z - z_0}{c}\right)^2,$$

where $leak$ is the aforementioned $leak$ parameter, eff is the combined efficiency of the MLP and the energy cut, the subscripts zero refer to the fitted position of the minimum (which is in reality very close to zero on the x and y axes, but at about +80 cm on the z axis), R_{min} is the minimal ^{210}Po rate, and a, b, c are shape parameters for the slopes of the parabolæ in the three Cartesian directions. Normalisation factors are omitted here for simplicity.

While Borexino's inherent symmetry causes the LPoF to form along the z axis (x=y=0), the properties of the positive temperature gradient, and other construction details, cause the LPoF to form above the equator. Additionally, effects like residual convection can cause the LPoF to move along the z-axis. While this movement is very slow (estimated maximum 20 cm per month), it makes it necessary to additionally separate our data in time-bins. In the following, we used a time-bin size of one month, but compatible results were achieved with larger bins. Even inside of the time bins, because of the very high number of points in the dataset, a three-dimensional binned likelihood method was further used in the spatial dimensions. This was necessary but also lead to a few problems of its own: it was found that this straightforward 3D fitting was too unstable, partly due to the relatively high number of parameters involved. To solve this, Eq. 19

was projected on a $(\rho^2 - z)$ plane:

$$\frac{d^2R(^{210}Po)}{d(\rho^2)dz} \propto (R_{min} * eff * geo + leak), \tag{20}$$
$$\text{with } geo = 1 + \left(\frac{\rho^2}{a}\right)^2 + \left(\frac{z - z_0}{b}\right)^2.$$

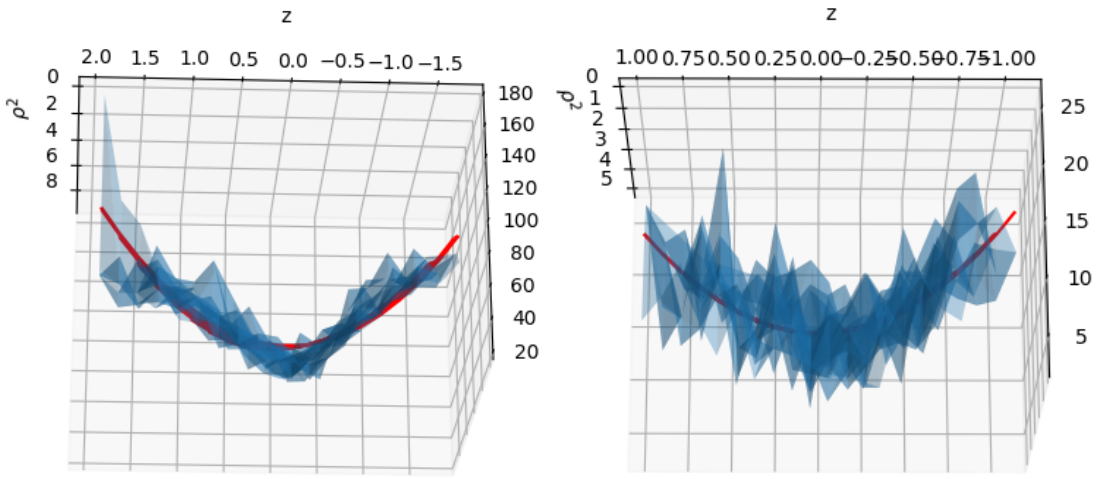
This of course assumes rotational symmetry along the z-axis. This symmetry is an expected consequence of Borexino's design, and was observed in tentative fits. While this approach decreases the available information through additional binning, it also reduces the number of free parameters to only four, namely R_{min}, a, b, and z_0, thereby solving stability issues.

Another critical point to consider is the complexity of the data further away from the minimum. The parabola estimation is only valid near the centre of the detector. When one approaches the nylon vessel, or other special points in the detector, the rate of ^{210}Po events (asymmetrically!) increases to levels that are orders of magnitude above the sought-after R_{min} value. In order to obtain sensible results, one must thus choose an appropriate fit region in which such edge effects, which would heavily disturb the fit, are negligible. The natural shape for such a region is an ellipsoid. Therefore, the following procedure was implemented in order to perform a fit in any given time bin:

1. Choose a mass M (the fit will be done in an ellipsoid encapsulating this LS mass).
2. Perform a grid search over shape/position parameters for that ellipsoid, but constrain it to always encompass exactly M (For the $\rho^2 - z$ fits this ellipsoid is symmetric along the z-axis and centered at x=y=0).
3. Choose the volume with the lowest mean rate of polonium.
4. Use Eq. 20 or Eq. 19 to fit the data in that volume (2D or 3D binned likelihood).

A first series of fits was performed in volumes of 100 t, one for every time bin. The outputs from these fits were then used as starting positions for fits in lower-mass volumes. This was repeated until the mass of the volumes reached about 20 t. The results from such fits can be seen in Figure 20. As one can see, this procedure allows one to fit the part of the data in which the paraboloid approximation is valid.

The z-position of the minimum of each fitted parabola, z_0, was then used to create a so-called *aligned* dataset. To create this dataset, the data was *blindly* aligned to the z_0 value of the previous month. This means that the data in each fitted time-bin i was shifted along the z-axis by the value z_0^{i-1} which was determined in the previous time interval.



(a) (b)

FIGURE 20: Example paraboloid fit results in a (a) 70 t volume and (b) 20 t volume. The binned data is represented in blue, and the fit results in red. The ρ^2 is in m^2, z in m, and the vertical axis is proportional to the rate of ^{210}Po. As one can see, the data in the larger volume is more complex than the model, but when going to lower masses, one can apply the fit function to the relevant data near the physical minimum.

This was done to avoid biases which can occur in minimum determination. Physically, this is also well motivated since the movement of the LPoF over one month is small relative to the size of the fitted paraboloid, as can be determined from a diffusion-only approximation using the Stokes-Einstein formula [224]. In general, it was deemed of the utmost importance that all sources of negative biases on R_{min} be avoided as much as possible. A negative bias in R_{min} would result in a positive bias on the rate of CNO neutrinos, thereby artificially boosting our signal significance.

While this method worked very well to a first order, some problems did arise. Firstly, the fit results were quite sensitive to the choice of the fit region. While this can be partly remedied by going to lower masses and using the procedure described above, it is clear that some effect remains, creating systematic uncertainties. Secondly, a period of high polonium activity in the lower detector early 2017 (see Figure 25), caused the fitted values of z_0 to jump up by almost one meter, in the affected time-bins, because the forced symmetry of Eq. 20 latched on to high rates on one side, thereby tilting the entire parabola. While this could in theory be resolved by further reducing the size of the fitted paraboloid, in practice the fits in volumes of less than about 20 t quickly become infeasible because of the low statistics. Thirdly, during the so-called ,,disruption" period in the second half of 2019, the fit became too unstable to use *i.e.*, it gave a value for R_{min} which, physically, could not correspond to the real minimum. However, the R_{min} values which were obtained in this period were always larger than the values obtained in other periods. This can easily be explained since during this disruption, ^{210}Po from

further away essentially covered the LPoF. Therefore, these abnormal values of R_{min} cannot bias the final result negatively and can still be used.

In principle, one could avoid using time bins in which the z_0 determination is problematic, however the reduction in statistics was found to be more detrimental to the discovery power of the spectral fit (see Sec. 3.4) than the change in the values of R_{min}. While I felt that explaining this basic approach in this section was important, these problems, as well as potential biases explained below and the aforementioned stability issues, can anyway be solved when switching to a Bayesian framework. This will be explained in more detail in the next sections.

3.6.2 Bayesian inference

Because of the extensive use of Bayesian inference techniques and nomenclature in the next sections, this one provides a detailed overview of the theory behind them. According to Bayes's theorem:

$$P(\theta|D,M) = \frac{P(D|\theta,M)P(\theta|M)}{P(D|M)}, \tag{21}$$

where θ represents the parameter array of the model M and D represents the available data. It follows that $P(\theta|D,M)$ is the *posterior* probability distribution of the parameters θ. $P(D|\theta,M)$ is the *likelihood* which can be calculated by comparing the data D with the results of the model M for any given array of θ. $P(\theta|M)$ is the usual Bayesian *prior* distribution of θ. Finally $P(D|M)$ is the Bayesian *evidence*:

$$P(D|M) = \int d^N\theta P(D|\theta,M)P(\theta,M), \tag{22}$$

which is a powerful tool to compare different models [225] that does not depend on any specific parameter array θ.

3.6.3 Evidence ratio

Because the evidence does not depend on θ, it is often ignored when calculating posterior distributions. However in a Bayesian framework, the evidence can be calculated for different models on the same data and used to determine which model is more likely to be true (Bayes factor analysis). Another way to write down Bayes's theorem in the context of model selection is:

$$P(M|D) = \frac{P(D|M)\cdot P(M)}{P(D)}. \tag{23}$$

In other words, the probability of a model given the data is equal to the probability of the data given the model (likelihood), multiplied by two terms that quantify prior beliefs about the model and the data. The Bayes factor (BF) [226], which is commonly used in model selection, can be written as:

$$\mathrm{BF} = \frac{P(M_1|D)}{P(M_2|D)} \cdot \frac{P(M_2)}{P(M_1)}. \tag{24}$$

In this example, the model M_1 is being compared to the model M_2 via the data D. In a Bayesian framework, the value of BF can thus be seen as the prior ratio, i.e. how likely the models are thought to be prior to seeing any data, multiplied by the posterior ratio, i.e. the likelihood ratio calculated from data. Thus, if the prior model probabilities are equal, then the BF is just a likelihood ratio. However, contrary to a so-called "likelihood-ratio test" (see [227] for great applications in physics), Bayesian model comparison does not depend on the specific set of parameters that maximises the likelihood but rather integrates over all possible parameter values. This expression of the Bayes factor can however lead to some problems. For example, the second term on the right-hand-side, the prior ratio, can be open to interpretation. While for ,,simple" applications this ratio is often one, for more complicated models more considerations can apply. However, this vagueness can disappear by simply inserting eq. (23) in eq. (24). One obtains:

$$BF = \frac{P(D|M_1)}{P(D|M_2)}. \tag{25}$$

Thus we see that when applying Bayes's theorem, the BF simply becomes the ratio of the evidences, which can nowadays be efficiently calculated numerically. The possible vagueness surrounding the model prior disappears into the much more straightforward determination of the parameter priors, which can in most practical cases be chosen to be uninformative around a physical range.

In order to get a more intuitive understanding of the meaning of this ratio (eq. (25)), we can look at a simple example. Consider a model where all parameters are normally distributed. In this case:

$$P(D|M) = \int d\theta^N \mathcal{L}(D|\theta, M)\pi(\theta|M) \tag{26}$$

$$= \int d\theta^N \frac{1}{\sqrt{\det(2\pi C)}} e^{\frac{1}{2}(\vec{\mu}-\vec{\theta})^T C(\vec{\mu}-\vec{\theta})} \pi(\theta|M), \tag{27}$$

where $\vec{\mu}$ are the central values of the Gaussian distributions, and C is the inverse covariance matrix. For flat priors between θ_{min} and $\theta_{min} + \delta\theta$:

$$P(D|M) = \left(\prod_{j=1}^{N} \frac{1}{\delta\theta_j}\right) \int_{\theta_{min}}^{\theta_{min}+\delta\theta} d\theta^N \frac{1}{\sqrt{\det(2\pi C)}} e^{\frac{1}{2}(\vec{\mu}-\vec{\theta})^T C(\vec{\mu}-\vec{\theta})}. \tag{28}$$

Since we are looking at a normal distribution, we can use a Laplace approximation (as in [228]) to obtain:

$$BF = \frac{\left(\prod_{i=1}^{N_1} \frac{1}{\delta\theta_j^{(1)}}\right)}{\left(\prod_{i=1}^{N_2} \frac{1}{\delta\theta_j^{(2)}}\right)} \cdot \frac{\mathcal{L}_{max}^{(1)} \sqrt{det(2\pi C^{(1)})^{-1}}}{\mathcal{L}_{max}^{(2)} \sqrt{det(2\pi C^{(2)})^{-1}}}, \tag{29}$$

where the (1) and (2) denote the two different models, and $\mathcal{L}_{max}$ is the maximum likelihood value. As one can see, we have recovered the same form as in Eq. 24 with a prior ratio multiplied to a likelihood ratio. The former can be seen as the ratio of the prior volumes. This is a quantitative version of Occam's razor, as models with large numbers of parameters will automatically be penalised (smaller evidence), and a very important tool that can for example be used to prevent over-fitting. Then, in a first approximation, the second term is proportional to a maximum likelihood ratio, thus also incorporating this information.

3.6.4 Nested sampling

In the following, the MultiNest algorithm [229–231] was used to calculate our results. MultiNest is a widely-used importance nested sampling algorithm. ,,Nested Sampling" [232] is a Monte-Carlo technique used to estimate the Bayesian *evidence* while generating posterior samples as a by-product. Because the calculation of the *evidence* is generally very complicated, *e.g.* needs marginalising over many nuisance parameters, it is often not analytically solvable and/or computationally expensive. It can, however, be estimated using Monte-Carlo techniques. In this case, the algorithm starts with N points (θ_1, ..., θ_N) randomly sampled from the prior, then iteratively calculates the hyper-volume encompassed by the points in parameter space. This is also called the prior-mass. In each iteration, low-likelihood points are removed and then replaced with points of higher likelihood, sampled from the prior. In this way, one creates a series of nested volumes which approach the region of maximum likelihood. Because the likelihood is a monotonically decreasing function of the prior-mass, one can then use eq. (22) in combination with the sampled likelihoods to estimate the value of the integral, therefore calculating the evidence. As a by-product of the algorithm, the remaining N sampled points at

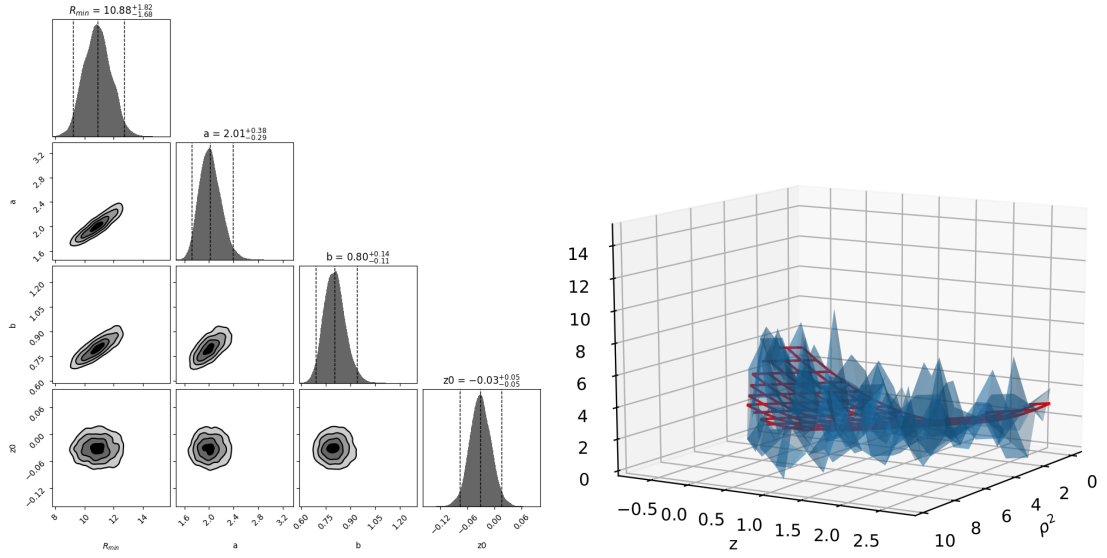
the end can be seen as samples from the posterior distributions of the parameters, as is common in Markov-Chain Monte-Carlo methods. Marginalising the points over any one of the parameters results in obtaining those distributions. MultiNest is a very efficient implementation of this algorithm, which further uses the discarded points in its estimation of the evidence. This, with other computational details, allows it to reach an order of magnitude higher accuracy and a greater stability than "traditional" sampling algorithms.

3.6.5 Bayesian approach

The work in the following sections was done in close cooperation with Sindhujha Kumaran in our working group. As an attempt to solve the previously mentioned difficulties in this measurement, a new Bayesian framework using MultiNest (see Sec. 3.6.4) was developed. A Bayesian approach has clear advantages in fits such as the one described in the previous sections, where many of the free parameters have clear physical interpretation and boundaries. For example, since we are looking for a paraboloid shape whose minimum leads to a rate measurement, it is clear that the parameters R_{min}, a, and b from Eq. 20 must be positive. While this can lead to problematic interpretations in a frequentist analysis, it can easily be dealt within a Bayesian framework. Additionally, since the Bayesian evidence (see Sec. 3.6.2) provides a very powerful tool to test different models, our framework can be used to quickly make conclusions about the current state of the data, *e.g.* whether it conforms more closely to a paraboloid or if a flat region already exists.

While other likelihood functions were used to investigate different effects, in the following, the same binned likelihood approach as in the last section was applied. In more detail: in each bin, the raw polonium hits N were counted and compared to model predictions M(θ) using Poisson statistics. For any parameter array $\vec{\theta}$, the value of the log-likelihood is therefore: $\mathcal{L} = \sum_{i=0}^{N_{bin}} -M(\theta)_i + ln(M(\theta)_i) \cdot N_i - ln(N_i!)$. For practical reasons, Sterling's approximation (to the second order) was used to calculate $ln(N!)$ in bins where $N > 15$. The results of typical fits can be seen in Figure 21. As one can see, the model is able to produce results that agree well with the data.

Additionally, because of the high stability of MultiNest in high-dimensional problems, the framework was also able to reliably fit the data using the three-dimensional fit (Eq. 19). The results of the three-dimensional fits were found to be compatible with the two-dimensional approach. Because the results of the latter are much easier to visualise, in the following, only results of the two-dimensional fits are presented. While most of the problems described at the end of the last section remain, the implementation of the



(a) (b)

FIGURE 21: Typical results of a fit using our Bayesian framework. (a) shows the posterior sample distributions of the different parameters, as well as correlation plots. As one can see, R_{min} is highly correlated with a and b, but independent of z_0. In (b), the data in blue is plotted with the fit results in red. The vertical axis is proportional to the rate of polonium.

Bayesian approach leads to a deeper understanding of the data, mostly from the ability to quickly test different models using Bayesian factor analyses, and from the results of the three-dimensional fits. This was also meant as a cross-check of the paraboloid approach, which was previously implemented by another working group in the Borexino collaboration.

While the Bayesian approach itself works quite well, the clear correlation between R_{min} and the shape parameters is also very visible (see fig. 21), underlying the importance of limiting the fit region to data just around the minimum, unaffected by residual convection and edge effects. However, since R_{min} is so affected by the data in the edges of the fit region, the choice of the region becomes a large source of systematic uncertainties.

3.6.6 Cubic spline extension

Another great advantage of the framework is that the optimisation process does not require knowledge of gradients, instead relying on repeated sampling of the likelihood. This allows one to easily use more arbitrary functions of the likelihood, and even algorithmic implementations. In order to solve the problems described in the previous sections, and to provide a more general solution to the paraboloid analysis, we implemented an extension of the Bayesian framework that uses cubic splines to fit the data. A cubic spline is a piece-wise polynomial of third degree. It is expressed through so-called

„knots" k which are simply points in the (x, y) space. The space between two knots is filled with a polynomial. Boundary conditions are then set: the polynomials on either side of one knot must both go through the knot. Their first and second derivatives must be continuous as well (exceptions exist for external knots). This results in a very smooth somewhat sine-like function which goes through each knot, see Figure 22(a). The high number of boundary conditions results in the function being completely and uniquely defined by the (x_k, y_k) positions of its knots, provided one has at least four knots with unique positions on the x axis. In general, any smooth function can be approximated to an arbitrarily high degree with a spline by increasing the number of knots. With low numbers of knots, one obtains parabola-like shapes which are characteristic of low-order polynomials.

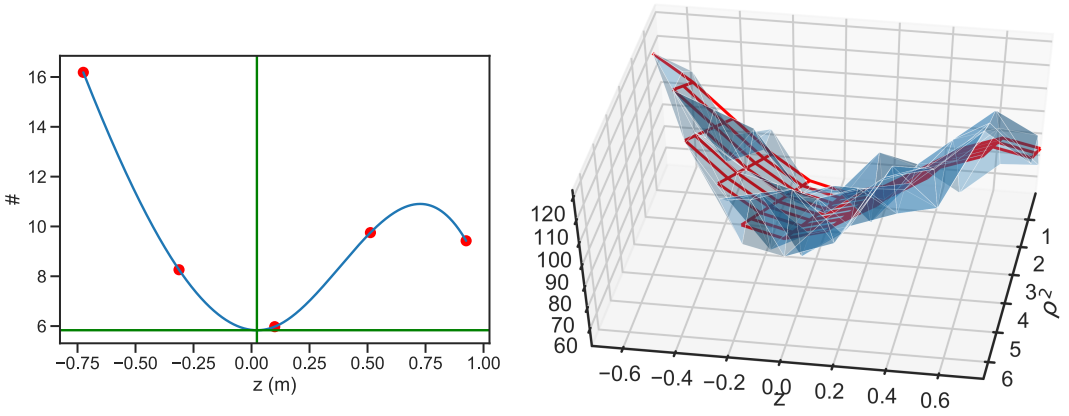
Because of the complex dependence of the derivatives of a spline on the position of the knots, it is very difficult to use a direct regression approach y = f(x). While it is possible to use a "standard" least-squares approach, as shown by Neitzel et al. (2019) [233] who implemented an iteratively linearized Gauss-Helmert model, the question of how many knots to use remains a difficult one to answer in a frequentist framework. However, both the methodology of the spline fit and the determination of an appropriate number of knots were in fact very easy to implement in our Bayesian framework. Because of the commonplace usage of cubic splines in image processing, spline functions are also already implemented and easily accessible in most programming languages[11], making it almost trivial to implement in our framework.

In the following, we first describe in details how a spline fit is performed with a given number of knots, and the next section explains how we determined the appropriate number of knots to use to fit the data. In order to perform a spline fit, Eq. 30 was used:

$$\frac{d^2R(^{210}Po)}{d(\rho^2)dz} \propto 1 + \left(\frac{\rho^2}{a}\right)^2 + spline(z). \tag{30}$$

This equation is analogous to Eq. 20, and can easily be extended to the equivalent 3D-form of Eq. 19. Normalisation and efficiency factors are omitted for simplicity. Here, the *spline* term refers to the spline function, which is intended to model the shape of the LPoF along the z-axis. Given k knots, the fit region is firstly determined from an ellipsoid, following the approach in Sec. 3.6.1. Then, k points are chosen uniformly on the z-axis, between the maximum and minimum z-bounds of the encompassing ellipsoid. These are the z-coordinates of the knots. The free parameters in the fit are the heights of those knots (proportional to the rate of ^{210}Po decays), as well as the shape parameter a. In this analysis, uninformative flat priors, limited by the physical boundaries when

[11] Good example: SciPy CubicSpline implementation



(a) (b)

FIGURE 22: Typical results of a spline fit using our Bayesian framework. (a) shows a cubic spline. The spline is completely defined by the positions of the knots, depicted as red points. The spline, in blue, is the unique solution that passes through all points while satisfying boundary conditions regarding second derivatives. The green lines are placed at the determined positions of $\propto R_{min}$ and z_0. In (b), the data in blue is plotted with the fit results in red in a grid generated using the spline in (a). The vertical axis is proportional to the rate of polonium, the z-axis is in (m) and the ρ^2-axis in (m^2). As one can see, the model agrees with the data very well in this 20-ton ellipsoid.

needed, were used for all parameters. Given the absence of R_{min} and z_0 parameters in Eq. 30, compared to previous methods, they must be subsequently determined from the spline itself, as explained below. Because the minimum of the ρ-term in Eq. 30 is zero, they can be determined independently of a. Figure 22 shows the results of a typical spline-fit. As one can see, they can be used to obtain a very good agreement with data even when the data is asymmetrically distributed along the z-axis, and in larger fit regions. We further observed that in the case of spline fits none of the parameters were correlated, except for some light correlation with a (see appendix B). Of course, the subsequent determination of R_{min} will still be correlated to spline parameters. However, while it will be proportional to the spline parameters whose z-coordinates are close to z_0, the spline parameters of far-away points will not affect it at all. This is the main strength of the approach.

Different methods can be used in order to determine R_{min} and z_0. Since MultiNest saves all sampled points of the likelihood points that it used during the nested sampling, we found that the fastest and most reliable approach was to directly use those existing points. For each point i, we calculated a corresponding R^i_{min} using the lowest point on the spline-curve (determined uniquely from each set of fit parameters) which also had a null first and second derivative. An example of such a minimal point can be seen as the green lines in Figure 22(a). To determine the total R_{min}, we weighed those values with the values of the likelihood l_i of the respective fits: $R^{weighed}_{min} = \frac{\sum_i R^i_{min} \cdot l_i}{\sum_i l_i}$. The same method was applied to determine z_0.

Measuring the complexity of the data As explained in the previous section, one can in principle fit arbitrarily complex functions using splines. Therefore, the agreement between fit and data can be tuned *ad absurdum* to infinitesimal disagreement by increasing the number of knots. To avoid this, and to find the balance between too few knots, which would lead to the loss of information, and too many, a Bayesian factor analysis (see Sec. 3.6.3) was used. To find the appropriate number of knots (i.e. to measure the complexity of the data along the z-axis) the following procedure was used:

1. Perform the paraboloid fit, note the evidence $\mathcal{Z}_0$.
2. Perform a spline fit with $k = 4$ knots, note the evidence $\mathcal{Z}_k$.
3. If $\mathcal{Z}_k \leq \mathcal{Z}_0$, use the results of the paraboloid fit.
4. Else, perform a spline fit with k+1 knots, note the evidence $\mathcal{Z}_{k+1}$.
5. Continue until $\mathcal{Z}_n \leq \mathcal{Z}_{n-1}$, then use $n - 1$ knots as final.

Here, special care had to be given in order to reliably compare evidence values. When using the spline model, where most parameters are uncorrelated, the posterior distributions for all parameters were always observed to be normally distributed, except for fringe cases where the size of the fit region was completely inappropriate. When using the paraboloid model, some problems however arose. For example, in the case of asymmetric data (e.g. in the period of early 2017 when the lower half of the detector saw a light surge in ^{210}Po decays), the posterior distribution of the shape parameter along the z-axis would usually peak at the slope of the parabolic shape directly around the physical z_0 minimum (i.e. around R_{min}), but another smaller peak could also be observed. This smaller peak corresponded to solutions which fitted the data further away from the physical z_0, and did not result in correct R_{min} values.

In theory, in our Bayesian framework, one could simply set the priors of the model parameters to be constant from 0 to infinity. In practice however, this is obviously not possible. Considering the aforementioned non-physical regions of parameter space, it was decided to follow an approach which ignored them while including all the necessary values in the regions of high likelihood. Empirically, we found that the we could perform a first fit with very broad priors, then a second one around the mode of each previously obtained posterior whereby the width of the priors was set to four (in both directions) standard deviations of said mode. In the case of normally distributed likelihoods and since the evidence is the integral of the likelihood over the prior, this allowed us to consider $\approx 99.994\%$ of all relevant likelihood points, while completely discarding non-physical regions which would artificially increase the evidence. This empirical rule was

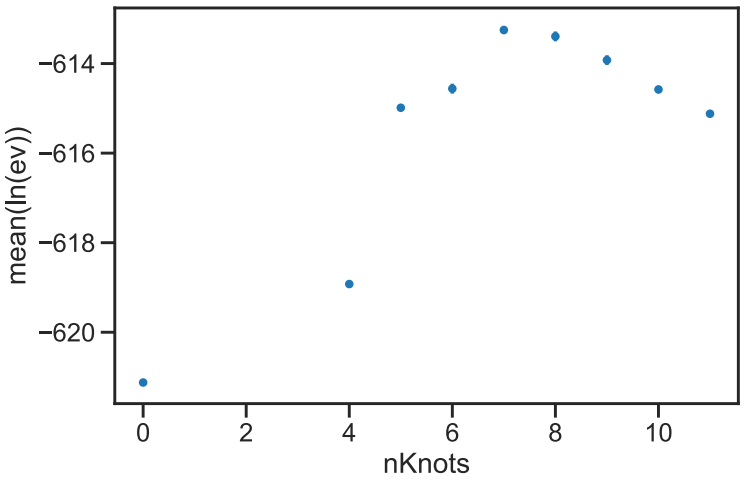
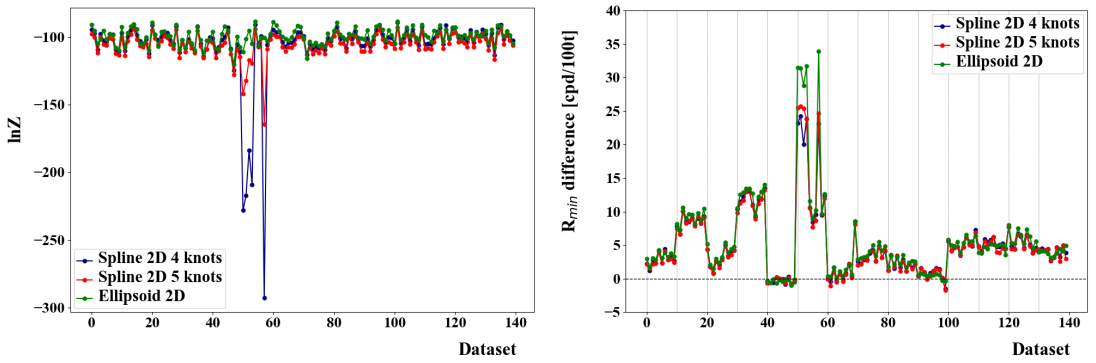


FIGURE 23: Average evidence value (over time) as a function of the number of knots per spline. The fits were performed in 100 t volumes. Zero knots corresponds to a paraboloid fit. As one can see, the evidence starts by going up with an increasing number of knots, then starts to go down again after a peak at seven. This can be understood as the point where the additional parameter cannot compensate in the likelihood what it costs in the "Occam" term. One can therefore conclude that the complexity of the (average) data best corresponds to that of a seven-knot spline. This plot nicely demonstrates the mechanism that prevents us from overfitting.

found to be valid in all time-bins considered in the final analysis. In repeated testing this approach also yielded very consistent values of the evidence. The mean values of the evidence over time, as a function of the total number of knots, can be seen in Figure 23. As one can see, we found that on average, a spline with seven knots is the best way to describe z-dependence of the data. Intuitively, as was observed on the data, one can think of a four-knot spline as very close to a parabola but with some permitted asymmetry. At five knots, one can start to model deviations from the parabolic structure in the points furthest away from the minimum. This might be why the jump in the evidence values between four and five knots is so high. With six and seven knots, some complexity is added to the deviations.

In order to test for possible sources of bias from our method, we validated the results against toy-MC simulations. The simulations, from another group in the collaboration, assumed a perfectly paraboloid LPoF with rotational symmetry along the z axis, and applied a range of different scenarios of convective motion in the detector using numerical fluid dynamics. The results of this validation can be seen in Figure 24. As one can see from the left plot, when applying our method to the toy dataset, the evidence values were in clear favour of the paraboloid approach (which corresponded to the underlying data). Additionally from the right plot one can see that in both cases, the results for R_{min} were greater than or equal to the real injected R_{min}. It is to be noted that the injected value of R_{min} was not known to us when the fits were performed. In general,

we concluded that in the toy-MC example, our method showed no bias toward negative R_{min} values. Additionally, it confirmed the validity of the procedure used to choose the number of knots, at least in the sense that it did not show preference for more complex fit functions when the data followed a simple paraboloid.



(a) (b)

FIGURE 24: Results of the validation of our methods against a toy-MC. In both plots, the x-axis corresponds to the dataset-id of a particular simulation. The ids are separated in batches of different physical conditions in the detector. The green points show the results of the paraboloid approach, and the blue and red points show the results obtained with a spline approach using four and five knots, respectively. As one can see in (a), the evidence (plotted logarithmically on the y-axis) is preferable for the paraboloid hypothesis, which in this case corresponded to the underlying data distribution. In (b) the difference between the injected and reconstructed values of R_{min} are plotted. While in some (physically improbable) scenarios, R_{min} was reconstructed with very high values, it was never reconstructed with significantly lower values.

3.6.7 Results on the ^{210}Bi constraint

The results of the methods discussed in the previous sections can be seen in Figure 25. In each time-bin, the data were fitted using the aforementioned procedures. The spline fits were performed only once and in volumes of 100 t, since it was not necessary to use the successive fits to lower masses, as is the case for the paraboloid method. The number of knots for a given spline was chosen to maximise the Bayesian evidence. That number was usually around six in stable regions (e.g. 2018) and around eight otherwise. In the region of disruption in 2019, the value of z_0 had to be chosen manually between two minima visible in the data. These can be understood as a previously clean region which was pushed down to negative z, and a new clean region forming again near $z_0 = 80$ cm. However, since we found that this choice had only insignificant influence on the final R_{min}, it is not discussed further. As mentioned in Sec. 3.6.1, the individual values of z_0 were then used to *blindly* align the data to avoid negative biases which are intrinsic to searches for a minimum. This final aligned dataset was then fitted using five knots (in a 22 t volume, see below), resulting in a value of $R_{min} = (11.15 \pm 0.83)$ cpd/100t for

the spline method and (12.04 ± 0.93) cpd/100t for the paraboloid fit. The value for the spline method was obtained from a dataset aligned using spline fits, and conversely for the result of the paraboloid approach.

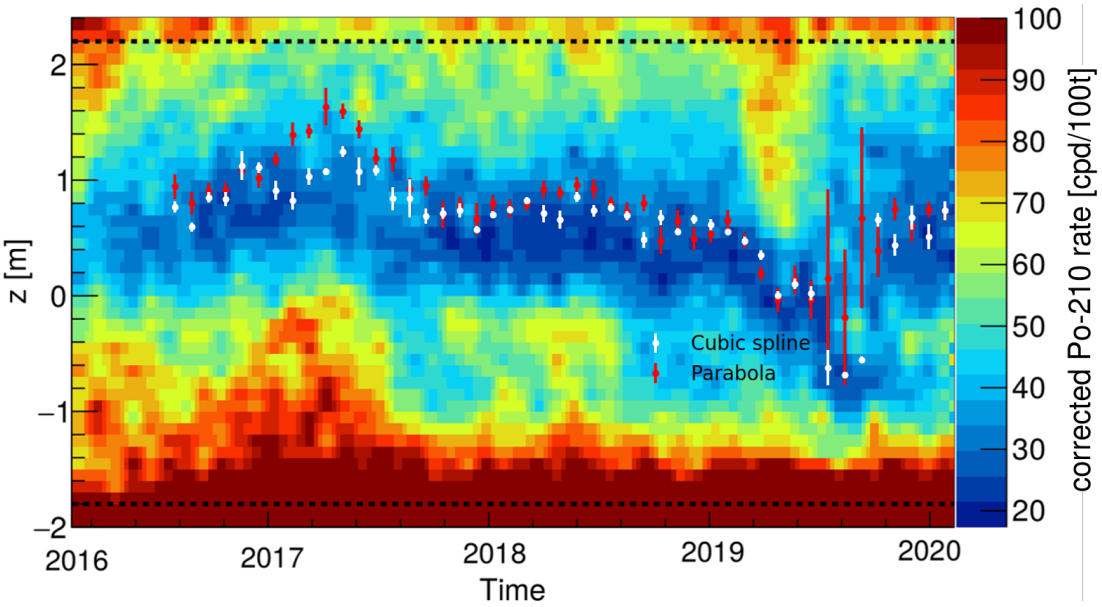


FIGURE 25: The time evolution of ^{210}Po decays in different z-slices of Borexino. Here, one can clearly see the LPoF: a region of low-activity moving around $z = 80$ cm. The red points represent the fitted positions of z_0 obtained from the Bayesian paraboloid fit, and the white points the z_0 positions obtained with the cubic spline fit. As one can see, the spline fit produces similar answers in stable regions, and more consistent answers in agitated periods. Movements of the region can be seen in early 2017, with an increased polonium activity in the lower half of the detector. The large bump in mid 2019 is due to human error that caused a turbulence which greatly affected the LPoF. However, after about half a year, the additional ^{210}Po (half-life of 138 days) which was brought into the LPoF has mostly decayed away, restoring the clean region.

The mass range used for this fit was determined by looking at p-values of the paraboloid fit. Indeed, we observed that p-values for low-masses were quite good (*i.e.* close to 0.3-0.5, but dropped as we increased the size of the fit region. This is of course because of the asymmetry of the data that cannot be modelled by the paraboloid model. Since we want to include as much of the data as possible, that can still be modelled, we chose to perform the fit on the aligned dataset in the largest volume in which the p-value was above 0.05, which was 15 t. The number of knots used in the spline-fit on the aligned dataset was determined by maximising the evidence in the volume of the same size.

In order to calculate the systematics, different effects were considered. Firstly, since we used a binned likelihood approach, that variations in the size of the bins have an influence on the final result. This influence can be seen in Figure 27(a). We varied the width of the bins from 10 cm to 29 cm (because outside of that range the behaviour of the fit was unstable). Secondly, the variation of R_{min} against the size of the fit region can be seen in Figure 27(b). To decide which mass range to use for this calculation, we

defined a *normalised likelihood*: the maximum likelihood (obtained from the sampling) divided by the total number of bins. Since we observed a jump of these values below 10 t and above 34 t, the systematic uncertainty was calculated in that range. Further, the central value of R_{min} quoted above was found by fitting in a 22 t ellipsoid (with a bin-width of 15 cm), because the *normalised likelihood* showed a clear minimum there (see Figure 26). The same was done for the paraboloid fit.

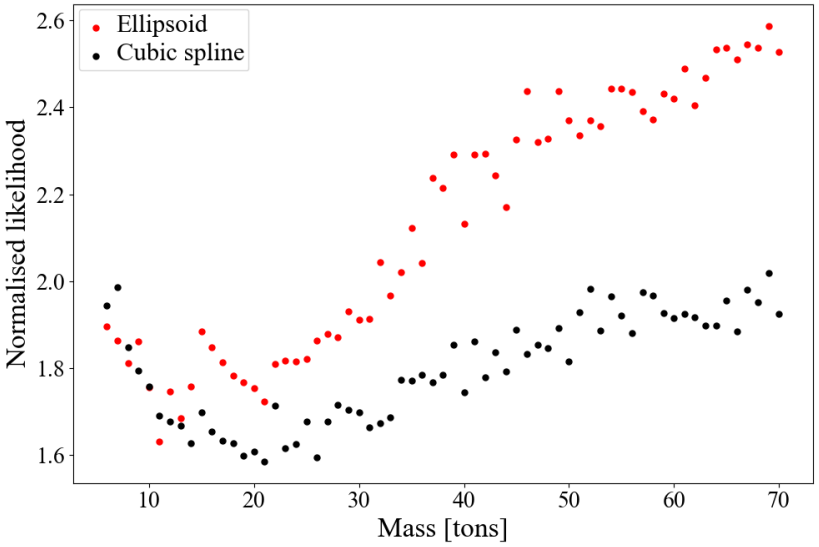
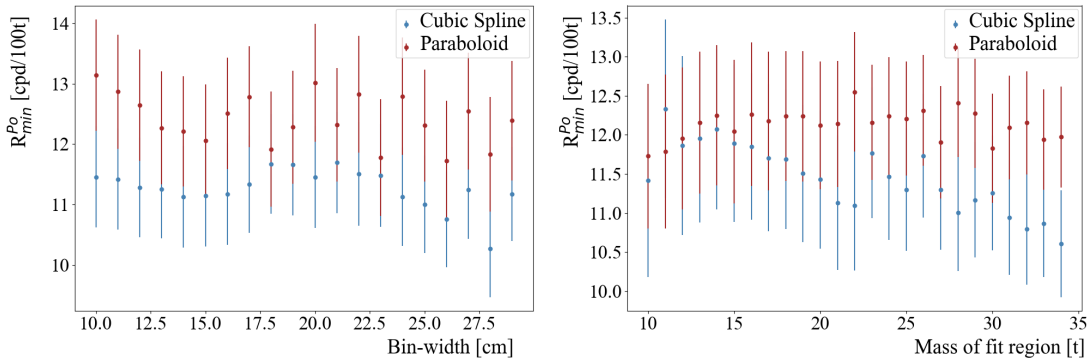


FIGURE 26: The fit of the (spline-)aligned dataset was done using five knots. The black points in this plot were calculated using a five-knot spline, and the red points using the paraboloid model. The y-axis shows the normalised likelihood, which is the maximum likelihood divided by the number of bins used by the fit, and the x-axis shows the mass of the LS encompassed by the fit region. As one can see, the normalised likelihood for the spline fit first drops, due to very low masses not having enough information for a meaningful fit, and then plateaus before rising again, due to the complexity of the data starting to exceed the five-knot hypothesis. The calculation of the systematics was performed in the plateau region where the model describes the data. While the red points are in near-perfect agreement with the black ones in very low masses, where the data is simpler than what the models describe, the trend of the red points for higher masses is similar albeit more extreme than the black ones, due to a less flexible model.

The third contribution originates from bismuth homogeneity calculations [68], the largest systematic uncertainty contribution to our result. The ^{210}Bi constraint from this analysis can only be used as an upper limit in the spectral fit if ^{210}Bi is homogeneously distributed over the fiducial volume ($r < 2.8\,\mathrm{m}$ && $-1.8\,\mathrm{m} < z < 2.2\,\mathrm{m}$). To calculate this homogeneity, all β events in an energy window with a maximal signal to noise ratio for ^{210}Bi were selected (15% of bismuth expected in the sample). The angular power spectrum, as well as the r^3 distribution of those events, can be seen in Extended Data Fig. 7 of [68]. Using an approach similar to studies of the cosmic microwave background, the figure shows that there is no evidence for a sizeable non-uniformity of β-like events



(a) (b)

FIGURE 27: The figures show calculated R_{min} values for different values of (a) Bin-width (in cm) and (b) fit region mass (in tons). The errorbars correspond to the statistical uncertainty on R_{min}, and the variation in the central values was used in the analysis to estimate systematic uncertainties. As one can see, the systematic uncertainties from those effects are relatively small compared to the statistical effects.

in the fiducial volume. When combining the uncertainties from both plots, one obtains an associated systematic uncertainty of 0.78 cpd/100t on the ^{210}Bi rate.

Finally, the uncertainty on the *leak* parameter introduced in section 3.5.1 was also considered as a source of systematic uncertainty (amounting to 0.3 cpd/100t).

In practice, two measures of systematic uncertainty are commonly used [234]. Firstly σ, the standard deviation of the points, which is interpreted as corresponding to a normally distributed measurement. Secondly, one often quotes $\sigma_\delta = \frac{max(x)-min(x)}{\sqrt{12}}$, which is the standard deviation expected from a uniform distribution. For further analysis of the results one might be interested in both. Interestingly, in the case of bin width and mass related systematics, we found that both definitions gave identical results for the spline systematics. The results using both definitions can be seen in Table 7.

Fit method	mass	R_{min}	Stat.	Sys.	Bi homog.	Total
Spline	22t	11.15	0.83	0.45	0.78	11.1 ± 1.3
Paraboloid(std)	15t	12.04	0.93	0.52	0.78	12.0 ± 1.3
Paraboloid(Δ)	15t	12.04	0.93	0.62	0.78	12.0 ± 1.4

TABLE 7: Results and uncertainties on R_{min} from different methods. The mass column indicated the size of the encompassing ellipsoid used for the final fit. std and Δ represent the two different methods of reporting systematic uncertainties (see text). When simply using the std definition, and combining results from both methods, one obtains $R_{min} = (11.5 \pm 1.3)$ cpd/100t. The „Sys." column includes all systematic effects except that from the bismuth inhomogeneity.

When combining the results from both the spline and paraboloid methods (with the σ definition of systematic uncertainty), the final ^{210}Bi constraint used in this analysis

becomes:

$$R(^{210}\text{Bi}) \leq 11.5 \pm 0.88^{stat} \pm 0.91^{syst} cpd/100t$$

3.7 The discovery of CNO neutrinos

The results from section 3.6 were published in Nature[12] [68] in which they were used to present the first experimental discovery of CNO neutrinos. This section is meant to give more details into the analysis and to discuss the significance of its results. Section 3.7.1 briefly describes a counting analysis, performed by another group in the collaboration, to give more insights in the results, section 3.7.2 explains the fit strategy used in the CNO analysis in detail, and section 3.7.3 presents the results. Finally section 3.7.4 discusses and concludes.

Events used in the CNO analysis were selected from the so-called *Phase-III* data spanning from July 2016 to January 2020 (after the thermal insulation campaign), with a total exposure of 7.64×10^4 d kt. Analysis cuts were applied to select events from a fiducial volume of 71.3 t situated in $r < 2.8$ m && -1.8 m $< z < 2.2$ m, and to reduce the rates of accidental muons, cosmogenic isotopes, other external backgrounds, electronic noise, and accidental coincidences. The selection cuts are $\approx 99.9\%$ efficient in selecting for neutrino events [137].

3.7.1 Counting experiment

A so-called Region-Of-Interest (ROI) can be chosen, defined as the energy region in which the contributions of CNO neutrinos, *pep* neutrinos, and ^{210}Bi represent at least 70% of the total event rate. One obtains an energy window spanning from 780 keV to 885 keV [68] in which the residual *background* consists mostly ($\approx 90\%$) of ^{7}Be and ^{11}C events.

A counting analysis was then performed where only the total number of events, from each contribution, was taken into account. The *pep* neutrino rate was constrained as explained in section 3.4 and the ^{210}Bi rate to (11.5 ± 1.3) cpd/100t. When further (conservatively) assuming the same ^{11}C rate as in Borexino Phase-II and sampling the ^{7}Be neutrino rate from HZ and LZ SSMs, the rate of CNO neutrinos was obtained by subtracting all components from the total measured number of events in the ROI by sampling from Gaussians with appropriate respective widths. This resulted in $R_{CNO} = (6.5 \pm 1.6)$ cpd/100t, on its own corresponding to a 3.6σ discovery of CNO neutrinos.

[12] https://www.nature.com/

The results can be further visualised on the left in Figure 28 as discussed in the following sections. The main conclusion of this analysis, which also serves as a cross-check, is that most of the sensitivity to CNO really comes from a small energy window.

3.7.2 Fit strategy

Since it is not possible to discriminate neutrino signals from background on an event-by-event basis, a multivariate fit was used in which the energy spectra of the events, split into a TFC-subtracted and a TFC-tagged spectrum (surviving or being rejected by a Three-Fold-Coincidence algorithm targeting ^{11}C), were fitted simultaneously, along with the radial distribution of the events.

The simultaneous fit to the three histograms was performed using a binned likelihood fit using a modified Poissonian approach. The likelihood had the form:

$$\mathcal{L} = \mathcal{L}_{tag}^{TFC} \cdot \mathcal{L}_{sub}^{TFC} \cdot \mathcal{L}_{RD}, \tag{31}$$

where the terms for the energy histograms were a Poisson likelihood:

$$\mathcal{L}_{TFC_X} = \prod_{i=1}^{N_X} \frac{\lambda_i^{n_i}(\vec{\theta})}{n_i!} e^{-\lambda_i(\vec{\theta})}, \tag{32}$$

where N_X is the number of bins in the corresponding histogram, $\lambda_i(\vec{\theta})$ is the number of events is the bin i, as predicted by the parameters $\vec{\theta}$, and n_i is the number of data events in the bin. The likelihood term for the radial distribution was:

$$\mathcal{L}_{RD} = \prod_{i=1}^{N_{RD}} \frac{a\lambda_i^{n_i}(\vec{\theta})}{n_i!} e^{-a\lambda_i(\vec{\theta})}, \tag{33}$$

where N_{RD} is the number of bins in the radial distribution histogram, and a is a scaling parameter designed to satisfy the following equation:

$$N = a \sum_{i=1}^{N_{RD}} \lambda(\vec{\theta}), \tag{34}$$

where N is the total number of events used to construct the radial distribution. Since the events in the radial distribution were obtained by projecting over an energy range included in the TFC-sub histograms, some events are counted more than once in the likelihood, inducing a correlation. The usage of the scaling term in eq. (34) is meant to reduce correlations by constraining the total number of events N to its known value.

Since only the Monte-Carlo fit approach was used (see Appendix A.4), all parameters in the (template) fit were integral rates corresponding to different spectral shapes, in the *nhits norm* variable, obtained from the simulation. The fit was performed in an energy range from 320 keV to 2640 keV (or 140 to 950 *nhits norm*), see appendix A.

The free parameters in the fit were the rates of CNO and ^{7}Be neutrinos, ^{85}Kr and ^{40}K decays in the scintillator, cosmogenic ^{11}C, gamma rays from external ^{40}K, ^{214}Bi, and ^{208}Tl. The rate of *pep* neutrinos was independently constrained (see Sec. 3.4). The ^{210}Bi constraint was enforced by using a half-Gaussian pull term in the likelihood function (flat below the ^{210}Bi constraint value, then Gaussian) to model the upper-limit determined in Section 3.6.7. In the end, the fit, which determined the rates of Monte-Carlo-produced shapes, yielded a p-value of 0.3, indicating a good agreement with the data. The best-fit curves can be seen in appendix B.

To evaluate systematics, the fit range and the binning were varied around the values used in the fit, and different theoretical shapes of ^{210}Bi decay were studied. In about 2500 fits, no significant changes in the rate of CNO neutrinos were observed. Then, as in [137], the parameters from the Monte-Carlo simulations were randomly varied in about 10.000 pseudo-experiments. The most important parameters were found to be, in that order: the spectral shape of ^{210}Bi, which can be constructed from different measurements [235–238]; the spectral shape of ^{11}C because of uncertainties in the energy scale of β^+ events; the non-linearity of the energy scale; the uncertainty in the light yield; and the inhomogeneity of the detector response. The random variations were constrained to values allowed by previously performed calibration measurements. This resulted in a systematic uncertainty of $^{+0.6}_{-0.5}$ cpd/100t on the rate of CNO neutrinos.

The resulting likelihood profile, with and without systematics, can be seen on the right in Figure 28, explained in more detail in the next section. When (quadratically) adding the systematic uncertainty to the statistical uncertainty of the fit, this results in:

$$R(CNO) = 7.2^{+3.0}_{-1.7} \text{ cpd/100t}.$$

3.7.3 Conclusion

As described in section 3.7.2, the rate of CNO neutrinos was determined to be $R(CNO) = 7.2^{+3.0}_{-1.7}$ cpd/100t using the multivariate fit. The large asymmetry on this result is largely due to the asymmetric nature of the ^{210}Bi constraint, which was implemented as a half-Gaussian pull term to model the lower limit. However, because the rate of ^{210}Bi was implemented as "only" an upper limit (translating into a lower limit for CNO) and not

a measurement, the resulting CNO value is also not to be interpreted directly as a measurement. In other words, one cannot directly interpret this number as a confidence interval on the true value of the rate of CNO neutrinos in Borexino, at least not without some Bayesian mathematics involved. However, the results of the fit can be used to quote a discovery power of CNO by trying to reject the hypothesis that no CNO was observed. From a relatively simple profile likelihood test (scanning the fit likelihood for different values of the CNO rate while minimising with respect to the other parameters), one obtains a significance of 5.1σ for the exclusion from $R(CNO) = 0$ hypothesis. In order to incorporate the effect of systematics inside of this estimation, we used Wilk's theorem to convert the profile likelihood curve to its corresponding log-likelihood with $ln\mathcal{L} = e^{-\Delta\chi^2/2.}$. This distribution was then convoluted with an asymmetric Gaussian with widths corresponding to the calculated systematic uncertainties ($^{+0.6}_{-0.5}$ cpd/100t), then converted back to a profile likelihood. The results can be seen in Figure 28.

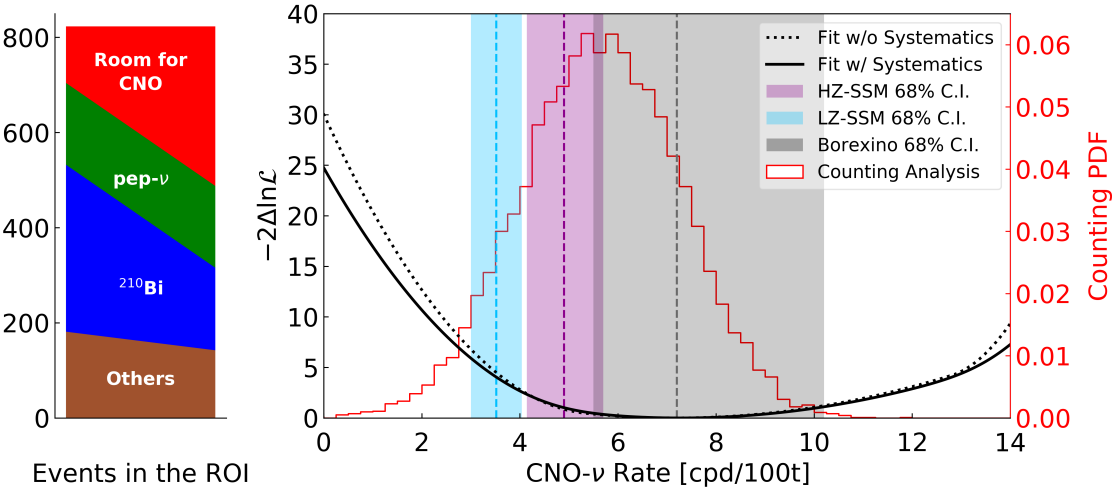


FIGURE 28: On the left, a visual representation of the counting analysis described in section 3.7.1, where one can see that even with the uncertainties, there is clear room for non-zero CNO. On the right, marginalised likelihood profiles from the fit (see section 3.7.2), plotted with a probability distribution sampled from the counting analysis. The solid and dashed black lines show the profiles with and without systematics, respectively. The strong asymmetric shapes are due to the ^{210}Bi lower limit, which was implemented as a one-sided Gaussian penalty in the likelihood calculation, but one can clearly see that while the systematic uncertainties slightly lower the significance of the results, one can still reject the null-hypothesis with 5σ (in the $\sigma = \sqrt{\Delta\chi^2}$ limit, see text for more detail). The red and blue shaded regions show the expected 68% CIs for the HZ and LZ SSMs, respectively.

Another more complex though exact method was used to calculate the significance with which one can reject the null hypothesis. This method used a likelihood ratio as test statistic [227]:

$$q = -2ln\frac{\mathcal{L}(CNO = 0)}{\mathcal{L}(CNO)}, \tag{35}$$

where $\mathcal{L}$ is the likelihood used in the fit, without and with the assumption of CNO, respectively. 13.8 million fits, performed on pseudo-experiments in which the spectral

shapes were deformed to estimate systematics (see section 3.7.2), and for the no-CNO scenario [68]. Using the so-obtained distributions as PDFs and the q obtained from data, a p-value corresponding to a significance of 5.0σ could be inferred at a 99% confidence level, see Figure 29. This is compatible with the result from the profile likelihood test.

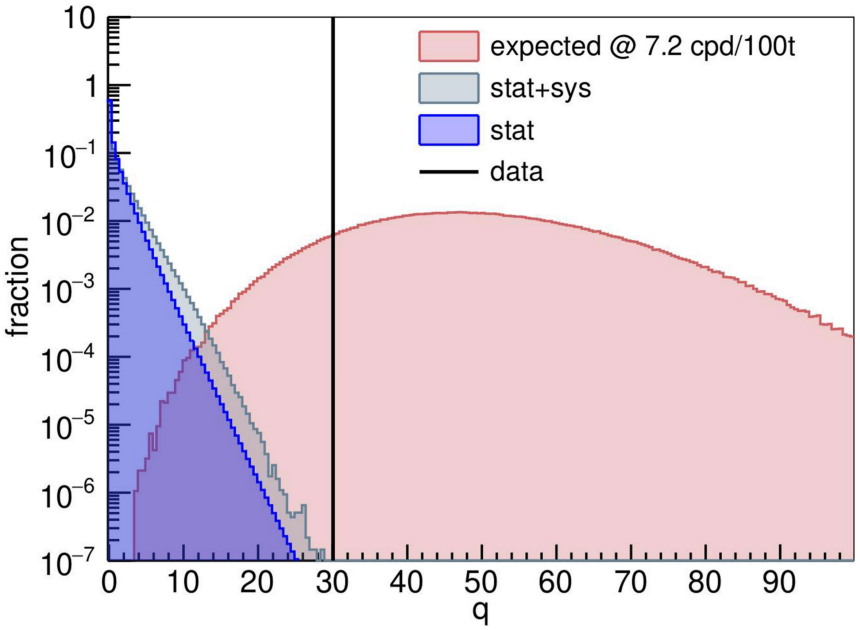


FIGURE 29: Distribution of the q_0 test-statistic (see text) in comparison. The blue histogram shows the distribution expected for the null hypothesis, when considering statistical uncertainties alone, as calculated using many toy experiments. The grey histogram is calculated in the same way but also including effects related to systematic uncertainties, as one can see, this slightly worsens the sensitivity. The red histogram is what is expected from toy experiments in which a rate of 7.2 cpd/100t was injected for CNO neutrinos. Finally, the black line shows the q_0 value measured from data at $q_0 = 30.05$, which corresponds to a 5σ discovery at the 99% percentile.

3.7.4 Outlook

The CNO analysis, as described in this entire chapter, might be improved in the future. More understanding of the low-energy region could allow the extension of the energy range down to about 190 keV (using pile-up in the fit), which would probably increase the overall sensitivity of the fit and allow the usage of the analytical fit approach. Further efforts are ongoing to try and fit the rate of CNO without using any external constraint on the rate of ^{210}Bi. This could be done by combining the Phase-III with Phase-II data, thereby using the boosted exposure (and therefore lower statistical uncertainties) to increase the discriminatory power of the fitter to the small but existing differences in the shapes of CNO neutrinos and ^{210}Bi decays.

Finally, especially since the 2020 pandemic caused significant delays in the schedule for the draining of the Borexino experiment, it might be possible to use ^{210}Po data

extending beyond January 2020 in order to improve the bounds on the ^{210}Bi constraint. Furthermore, in a longer period of thermal stability, we might obverse a plateau forming near the centre of the detector, and therefore claim not "only" a discovery but also a direct measurement of the rate of CNO neutrinos. The Bayesian model developed in section 3.6.6 is particularly powerful in this scenario, since its innate ability to use Bayesian evidence to compare different models can be used to infer whether a plateau-model is preferred over a more paraboloid-like shape.

In further hypothesis testing, when assuming the validity of the MSW-effect [218], neutrino oscillation parameters from [239], and an electron density of $(3.307 \pm 0.015) \cdot 10^{31}$ per 100t, the rate of CNO events in Borexino can be converted to a flux of CNO neutrinos of $7.0^{+3.0}_{-2.0} \cdot 10^{8}\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ on Earth. While this result is compatible with both HZ an LZ models of solar metallicity within 0.5σ and 1.3σ, respectively, when combined with previous measurements of solar neutrinos by Borexino [137] it unveils a favour for the HZ model at 2.1σ (compared to about 1.9σ without this measurement [46]).

While the expressiveness of the hereby quoted value of CNO, with an uncertainty of almost 25%, is limited in terms of direct interpretation (as expected from [60]), considering the high significance with which it rejects the null-hypothesis, we can confidently claim that the work presented in this section was used for the first direct detection of CNO neutrinos from the Sun. This paves the way for next-generation experiments to confirm and refine this measurement. As explored in section 1.3.3, these measurements will probably finally solve the long-lasting problem of solar metallicity, and lead us to a better understanding of the stars all around us.

4 JUNO solar neutrino sensitivity studies

> Why should we look to the past in order to prepare for the future? Because there is nowhere else to look.
>
> James Burke, *Connections*

4.1 Introduction

As stated in section 1.3, solar neutrinos not only represent a unique opportunity to study the deepest inner workings of the Sun, but also constitute a high intensity neutrino source that can be used to study neutrinos themselves. Historically, solar neutrinos have been used since the 1960s by radio-chemical experiments like Homestake [64] to help develop the standard solar model (see sections 1.1 and 1.3.1) and provide profound insights in the nuclear processing powering our star. They were also used to study neutrinos themselves, for example with measurements from Super-Kamiokande [12] and SNO [18] which provided a direct proof of neutrino flavour conversion. More recently, Borexino measured all pp-chain solar neutrinos [46], as well as CNO neutrinos [68]. The latter constituted the first direct evidence for those neutrinos as well as for the existence of the nuclear process in the Sun, and the work in this thesis (see section 3) directly contributed to it.

Due to its unprecedented size and sheer quantity of liquid scintillator, the JUNO detector (see section 2.2) will be able to measure an unrivalled amount of solar neutrinos through elastic scattering with an unprecedented energy resolution. With simple *back-of-the-envelope* estimations one can already show that JUNO has the potential to detect about 5000 ^{7}Be neutrinos per day in its fiducial volume, which is about 142 times more than in Borexino. One can therefore expect unprecedentedly precise solar neutrino measurements from JUNO within a few years or measurement. However, due to the technical difficulties surrounding the purification of liquid scintillator and the fact that JUNO's main goal is oriented towards reactor neutrinos, it is unclear what JUNO's sensitivity will be for solar neutrinos. As a reminder, it is much easier to reject backgrounds for reactor neutrinos through IBD detection (see section 1.6.5) than for solars through elastic scattering, which makes the detection of solars intrinsically more dependent on backgrounds. It is therefore particularly important to study JUNO's sensitivity as a function of different background levels. This section presents a sensitivity study of JUNO to *pep*, ^{7}Be, and CNO neutrinos in an energy range from 0.45 MeV to 1.7 MeV. I excluded *pp*

neutrinos from the analysis in this thesis because the difficulties related to the very low-energies are outside of its scope. These difficulties encompass the inclusion of (mostly ^{14}C-related) pile-up and necessary improvements of the very low energy parts of the simulation of the detector's trigger system. Sections 4.2 and 4.3 list the different signal and background contributions, respectively, section 4.4 presents the methods used for the analysis, and section 4.5 shows the sensitivity results.

4.2 Expected signal

Since solar neutrinos in JUNO are detected through elastic scattering, there is in principle no energy threshold. However in practice the huge amount of ^{14}C background does represent a limiting factor for the analysis - not least because the sheer number of expected ^{14}C events in JUNO make it impracticable to simulate. Indeed, the exact treatment of the trigger in the electronic simulation is at the time of writing lacking - which makes it beyond the scope of this thesis to correctly handle events at the lowest energies. The region of interest (ROI) in this analysis is therefore an energy window starting at 0.45 MeV and going up to 1.7 MeV (visible energy). This upper-limit was chosen to maximise the precision of the solar neutrino measurements. The expected event rates in JUNO and in the ROI can be seen in table 8.

Solar ν	R_{JUNO} (cpd/kton)		R_{ROI} (cpd/kton)	
	HZ	LZ	HZ	LZ
^{7}Be	490 ± 29	447 ± 26	142.5 ± 8.3	130.0 ± 7.5
pep	28.0 ± 0.4	28.4 ± 0.4	17.1 ± 0.2	17.3 ± 0.2
CNO	50.3 ± 8.0	36.0 ± 5.3	16.6 ± 2.6	11.9 ± 1.8

TABLE 8: Expected solar neutrino event rates in JUNO and in the ROI from 0.45 MeV to 1.7 MeV. The predictions are separated between the HZ (high metallicity) and LZ (low metallicity) SSMs (see section 1.3.1).

4.3 Expected backgrounds

For liquid scintillator solar neutrino analyses, the expected background is usually much larger than the signal events, which makes shape-based analyses an absolute necessity. In the following, I describe the four sources of background events: *internal*, referring to intrinsic scintillator radioactivity, *external*, referring to gamma rays penetrating into the scintillator from the outside, *cosmogenic*, which stem from the decay of isotopes that are themselves generated in the spallation of mostly ^{12}C atoms by cosmic muons entering the detector from above, and finally *pile-up*, which originates when events occur too close in time. This section is mainly based on work I performed with Davide Basilico and

Luca Pelicci, as documented in Ref. [240]. Other sources of background not cited below, such as backgrounds from the elastic scattering of reactor neutrinos, were estimated to be negligible and are therefore not considered in the analysis.

4.3.1 Internal radioactivity

^{85}Kr is a beta emitter with a mean lifetime of $\tau = 15.4\,\mathrm{yr}$ and an endpoint energy of 687 keV. The isotope is present in high quantities in the atmosphere, globally around 5000 PBq, originating mostly from nuclear fuel reprocessing [241]. Consequently, even small exposures to air can introduce relatively large contamination in the liquid scintillator. Due to its similar spectral shape to that of ^{7}Be neutrinos, it constitutes a particularly problematic background.

^{40}K can undergo beta decay with an endpoint energy of 1.31 MeV and a branching ratio of 89.3%, or a gamma decay with an energy of 1.46 MeV and a branching ratio of 10.7%. Its total mean lifetime is $\tau = 1.85\,\mathrm{Gyr}$. Its rate and the spectral shape of its beta spectrum also cause it to be a nefarious background *e.g.* for *pep*-neutrinos, but it can in general be constrained quite well through its gamma peak (though only if a working TFC algorithm is present because otherwise a large ^{11}C background is expected to cover these peaks, see section 4.3.3).

^{232}Th chain With a mean lifetime of $\tau = 20.3\,\mathrm{Gyr}$, ^{232}Th is a long-lived primordial nuclide that still permeates most of natural silicate rock formations. Its long decay chain encompasses six alpha decays and four beta decays - and for simplicity they are all summarised as ^{232}Th chain in this analysis. Along with ^{238}U decays, they constitute one of the main background sources in JUNO.

^{238}U chain Similarly to the ^{232}Th chain, ^{238}U emits eight alpha and six beta particles along its own decay chain. With a mean lifetime of $\tau = 6.45\,\mathrm{Gyr}$, it is also a primordial nuclide though its natural occurrence is slightly lower than that of ^{232}Th.

^{210}Pb sub-chain ^{210}Pb is one of the last elements in the decay chain of ^{238}U. It undergoes the following decay (sub-)chain:

$$^{210}\mathrm{Pb} \xrightarrow[\mathcal{Q}=63.5\,\mathrm{keV}]{\tau\approx 32\,\mathrm{y}} {}^{210}\mathrm{Bi} \xrightarrow[\mathcal{Q}=1160\,\mathrm{keV}]{\tau=7.23\,\mathrm{d}} {}^{210}\mathrm{Po} \xrightarrow[\mathcal{Q}=5.3\,\mathrm{MeV}]{\tau=199.1\,\mathrm{d}} {}^{206}\mathrm{Pb}, \tag{36}$$

While ^{210}Pb on its own is not problematic, its decay products ^{210}Bi and ^{210}Po are. Looking at the lifetimes of the isotopes, it is clear that any ^{210}Pb contamination would quickly generate ^{210}Bi decays, and after about two years the chain would be in equilibrium with ^{210}Po decays on the same level as the two others. This contribution would however be out-of-equilibrium with respect to the ^{238}U chain. Initial contaminations of ^{210}Pb in the liquid scintillator were observed by the Borexino collaboration and can be expected to occur in JUNO as well. In addition to the ^{210}Bi and ^{210}Po from the ^{238}U and ^{210}Pb chains, Borexino also observed *unsupported* ^{210}Po, probably originating from the pipes containing the scintillator as well as from ^{210}Pb that had deposited on Borexino's nylon vessel. This was a problem because the ^{210}Po could detach from the vessel and enter the scintillator (neither ^{210}Pb nor ^{210}Bi detachment was observed). While an *unsupported* ^{210}Po source would decay away relatively quickly, additional ^{210}Po from unexpected ^{210}Pb could become quite problematic.

^{14}C is the highest background in liquid scintillator. Being naturally abundant in organic matter to a level of 10^{-12} compared to ^{12}C atoms, it is mostly generated in cosmogenic reactions in the atmosphere and transmitted to plant (and thus animal) matter through photosynthesis. Thankfully, its low energy endpoint of 156 keV means that it doesn't affect large parts of the energy spectrum except for pp-neutrinos whose spectrum is in large parts hidden by ^{14}C. However, because of the high rates, pile-up of ^{14}C (with itself and other particles) is expected to possibly severely affect the analysis, see section 4.3.4.

4.3.2 External backgrounds

External backgrounds are gamma rays which originate outside of the liquid scintillator volume but penetrate into and end up depositing energy in the scintillator. These gamma rays mostly originate from ^{208}Tl, ^{214}Bi, and ^{40}K that are present in the PMT and its surrounding materials. JUNO's AV - with a diameter of 35.4 m, rests in a 43.5 m diameter and 44 m high cylindrical water pool which shields it from external backgrounds. Because of this, gamma rays with an energy below about 1 MeV cannot reach the LS. Further, preliminary Monte-Carlo simulations show that a radial fiducial volume cut in the LS of about 14 m to 15 m is enough to completely remove any external background, which is no problem given the huge mass of the JUNO target.

4.3.3 Cosmogenic isotopes

Cosmogenic isotopes are produced by the spallation of atmospheric muons on mostly ^{12}C atoms in the scintillator. Many of the thereby created isotopes are short lived and can be removed with a simple veto cut around the muon track, but some isotopes are somewhat long-lived and become so-called *cosmogenic* background. The relevant isotopes for this analysis are ^{10}C, ^{11}C, and ^{6}He, with half-lives of 27.8 s, 29.4 min, and 1.1 s, respectively. This is high when compared to the average time between muon events, which in JUNO is about one third of a second. The most important cosmogenic background for this analysis is by far ^{11}C, because of its high rates and because it covers large sections of the *pep* and CNO neutrino spectra. It can however be effectively tagged, at the cost of some exposure, using a threefold-coincidence (TFC) algorithm [242]. This algorithm divides the dataset in two: the TFC-tagged and TFC-sub datasets. TFC-tagged contains all events which were tagged as ^{11}C candidates by the TFC algorithm, and inversely. Because a TFC algorithm doesn't yet exist in JUNO, these histograms were created using assumptions on the expected performance of the future algorithm. Based on Borexino experience, a tagging power (TP) of 0.9 was assumed, and a subtracted exposure (SE) of 0.7. These values of TP and SE mean that the TFC-tagged dataset is expected to include 90% of the ^{11}C events but only 30% of the total exposure. This is a direct consequence of the veto-cut inside of the TFC algorithm that can efficiently tag ^{11}C evens but also includes large portions of the detector (in time **and** space).

The expected rates of cosmogenic backgrounds in JUNO can be estimated using measurements by other similar experiments. In this case, it is even necessary because, at the time of writing, the corresponding JUNO simulation software is not ready for such calculations. The scaling was done using the following formula:

$$R^{JUNO} = R^{X} \cdot \left(\frac{\bar{E}_{\mu}^{JUNO}}{\bar{E}_{\mu}^{X}} \right)^{\alpha} \cdot \frac{\Phi^{JUNO}}{\Phi^{X}} \cdot \frac{\epsilon_{C}^{JUNO}}{\epsilon_{C}^{X}}, \tag{37}$$

where X denotes the corresponding scaling experiment, R is the production rate in cpd/kton, $\bar{E}_{\mu}$ is the average muon energy at the experimental site, α is the spectral index of the energy dependence on the isotope production yield [243], Φ is the total muon flux at the site, and ϵ_C is the mass fraction of carbon in the scintillator. Since JUNO's corresponding analysis is not finalised, this formula simply assumes that the detection efficiencies of JUNO will be similar to that of the other experiments. The two experiments used to perform this scaling are Borexino [181] and KamLAND [167]. The experiment-related values used for the scalings can be found in table 9, and the isotope-related parameters in table 10. Finally, the values obtained from the scaling can be found in table 11.

X	$\bar{E}_\mu$ (GeV)	Φ ($m^{-2}\,h^{-1}$)	ϵ_C
JUNO	209.2 ± 6.4	10.8 ± 1.1	0.8792
KamLAND	260 ± 8	5.37 ± 0.41	0.8568
Borexino	283 ± 9	1.236 ± 0.001	0.9007

TABLE 9: Parameters used in the scaling of Borexino and KamLAND cosmogenic rates to JUNO in eq. (37). The numbers are from [92, 191, 243, 244] and were summarised in [240].

Isotope	R^{BX}	R^{KL}	α
^{11}C	274 ± 3.2	1106 ± 8	0.703 ± 0.002
^{10}C	6.2 ± 2.2	21.1 ± 1.8	0.810 ± 0.010
^{6}He	11.1 ± 4.5	15.4 ± 2	0.818 ± 0.013

TABLE 10: Rates of cosmogenic backgrounds in Borexino and KamLAND [191, 244] in units of cpd/kton, and the used values of the spectral index α [243].

Isotope	Experiment	R^{JUNO} (cpd/kton)	$\langle R^{JUNO} \rangle$ (cpd/kton)
^{11}C	Borexino	1890 ± 199	1916 ± 157
	KamLAND	1959 ± 254	
^{10}C	Borexino	41.4 ± 15.3	37.1 ± 5.3
	KamLAND	36.5 ± 5.7	
^{6}He	Borexino	74 ± 31	27.8 ± 4.8
	KamLAND	26.6 ± 4.9	

TABLE 11: Estimated cosmogenic background rates in JUNO based on scalings from the Borexino and KamLAND experiments using eq. (37). The right-most column shows the weighted average between both calculated scalings.

4.3.4 Pile-up

Pile-up occurs when two events are so close in time that they cannot be properly disentangled by the experiment's reconstruction methods. In the case of the liquid scintillator detection of elastic scatterings, this often simply means that two separate physical events were reconstructed as a single higher-energetic event. This is in theory quite rare, due to the nature of Poisson statistics, however, when dealing with large event rates, this can become problematic. Additionally, two events which were originally outside of the FV could, if mis-identified as a single event, be reconstructed as having occurred inside it. Interestingly, while it was found that in our energy region the pile-up of two ^{14}C is insignificant, the pile-up of four is not. While this seems to contradict Poisson statistics, it is actually a reflection of the amount of pile-up events that can reach our energy range: while two ^{14}C events don't have enough energy to reach the spectral range, four ^{14}C events do, even if the probability of "four ^{14}C pile-up" is much lower. The other significant source of pile-up occurs when a ^{14}C event mixes with a ^{210}Po event. Both of these components can be seen in fig. 30. Other possible pile-up components are ^{210}Po-^{210}Po and externals with any other component because of their broad energy spectrum

and very high rates, but their impact was found to be negligible. While the impact of ^{14}C-^{14}C and ^{14}C-^{210}Po pile-up was found to be negligible on *pep* and CNO, its impact on ^{7}Be is more nuanced, see section 4.5.1

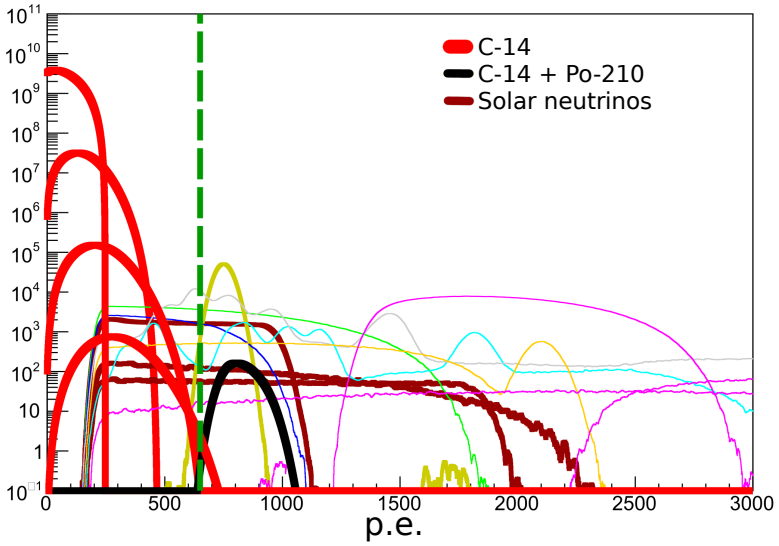


FIGURE 30: Energy spectra of different components (in p.e.) in one year of data-taking in the *baseline* scenario (see section 4.3.5). The ROI of this analysis starts at the dashed green line. The red components show the expected amount of ^{14}C in the detector in different stages of pile-up, up to quadruple (see text). The curves were generated analytically and are to be seen as purely representative. As one can however see the quadruple component is the only one which intersects, albeit only slightly, with the ROI. The black curve shows the expected ^{14}C-^{210}Po pile-up, which as one can see is much more problematic.

4.3.5 Scenarios

As will be seen below, JUNO's sensitivity to solar neutrinos is mainly dependent on the levels of radiopurity that will be achieved in the LS. JUNO's LS is based on linear alkylbenzene (LAB) as a solvent with an added 2.5 g l^{-1} of 2,5-diphenyloxazole (PPO) as fluor and 3 mg l^{-1} of p-1,4-bis(2-methylstyryl)benzene (bis-MSB) as a wavelength shifter. These concentrations were optimised using dedicated tests with a Daya Bay detector [193]. The background levels assumed in this thesis are based on different assumptions but are based on a lot of work by the JUNO collaboration to reduce the radioactivity in JUNO in general and inside of the LS [92, 203]. These assumptions on the contamination of the liquid scintillator can be converted to an expected number of event in JUNO with the following formula:

$$R_i = \frac{c_i \cdot ln(2)}{t_{1/2} \cdot m_i \cdot 10^3 \cdot M} \cdot 86400 \cdot 10^9, \tag{38}$$

where R_i is the event rate of isotope i in cpd/kton, c_i is its concentration in the LS in g/g, $t_{1/2}$ is its half-life in s, m_i is its molar mass in Da, M is the atomic mass constant in kg, and the other factors are for conversion. The assumptions on the LS contamination are summarised in table 12 and their effect can be seen in fig. 31. An example dataset based on those assumptions can be seen in fig. 32. In the following, four radiopurity scenarios are considered when looking at prospective sensitivity:

IBD This corresponds to the minimal requirements as set by the JUNO collaboration [92]. With ^{238}U and ^{232}Th concentrations of 10^{-15}, this is the maximum level of contamination that still allows for the mass hierarchy to be determined at the level of 3σ in six years of data-taking.

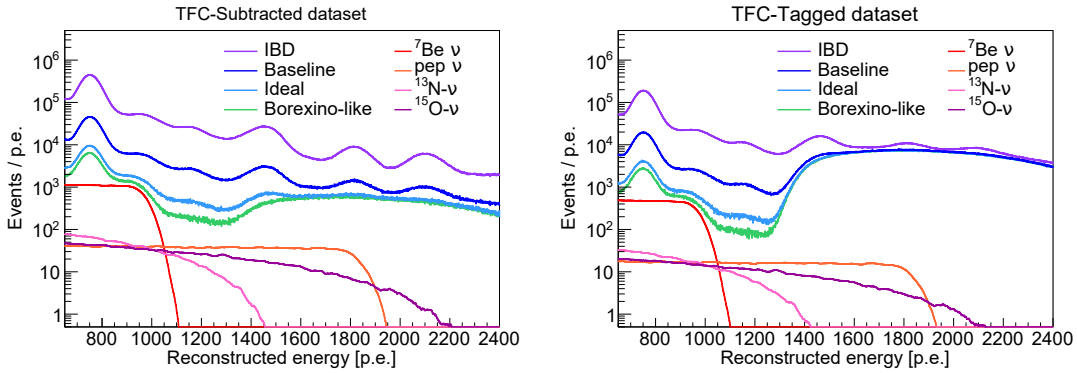
Baseline This is mostly identical to the *IBD scenario* but with a factor ten improvement of all internal backgrounds.

Ideal This corresponds to the other set of assumptions from Ref. [92]. The contamination from internals is again roughly ten times lower, except for ^{210}Pb and ^{85}Kr where the improvement is closer to a factor of five.

Borexino This scenario was only added as a reference since it is not expected to be reached in JUNO. It assumes identical contaminations for ^{238}U, ^{232}Th, ^{40}K, and 85Kr as were achieved in Phase-III of the Borexino experiment [110].

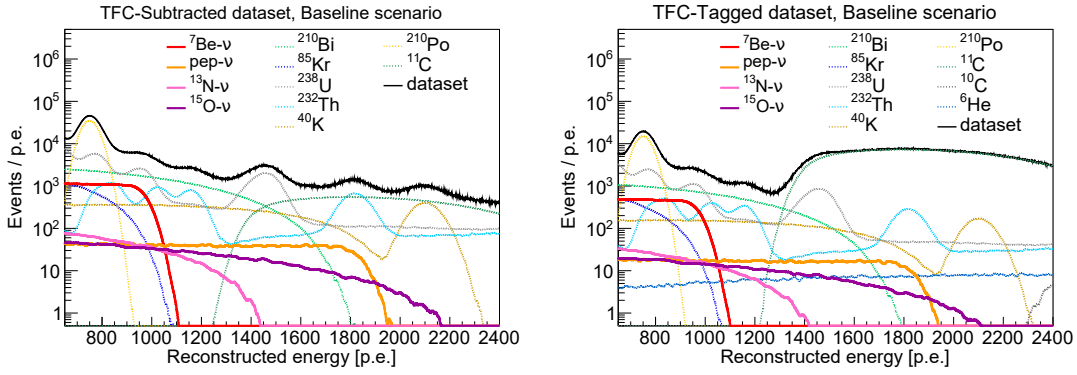
		^{40}K	^{85}Kr	^{232}Th chain	^{238}U chain	^{210}Pb chain
IBD	c	10^{-16}	$4 \cdot 10^{-24}$	10^{-15}	10^{-15}	$5 \cdot 10^{-23}$
	R	2289	5000	3508	15047	36817
	R^{ROI}	1562	705	2100	7368	17269
Baseline	c	10^{-17}	$4 \cdot 10^{-25}$	10^{-16}	10^{-16}	$5 \cdot 10^{-24}$
	R	229	500	351	1505	3682
	R^{ROI}	156	70	210	737	1727
Ideal	c	10^{-18}	$8 \cdot 10^{-26}$	10^{-17}	10^{-17}	10^{-24}
	R	23	100	35	150	736
	R^{ROI}	16	14	21	74	345
Borexino	c	$\leq 2 \cdot 10^{-19}$	$\leq 8 \cdot 10^{-26}$	$\leq 5.7 \cdot 10^{-19}$	$\leq 9.4 \cdot 10^{-19}$	$5 \cdot 10^{-25}$
	R	4.2	100	2	1.4	347
	R^{ROI}	2.9	14	1	1	163

TABLE 12: Internal background rate assumptions of different isotopes in different scenarios. c refers to the concentration of the isotope in the liquid scintillation in g/g, R is the corresponding event rate in cpd/kton, and R^{ROI} is the event rate inside of the ROI (see text).



(a) (b)

FIGURE 31: Expected neutrino spectra in one year of data-taking for the (a) TFC-Sub and (b) TFC-Tag histograms. The background contributions for the four different scenarios are also shown for comparison. The curves were generated assuming TP=0.9 and SE=0.7 (see text).



(a) (b)

FIGURE 32: Expected spectra in the *baseline* scenario in the (a) TFC-subtracted and (b) TFC-tagged dataset. The neutrino components are shown as bold lines, and an example expected dataset as the black connected points. One can clearly see the difference in the levels of ^{11}C in both plots, as well as the difference in the overall scales - both stemming from the TFC algorithm performance. The curves were generated assuming TP=0.9 and SE=0.7 (see text).

4.4 Analysis

This section presents the methods and software used to calculate JUNO's sensitivity to ^{7}Be, *pep*, and CNO solar neutrinos. Section 4.4.1 explains the principles behind the sensitivity calculations and section 4.4.2 elaborates on the software which was developed in this context - which was designed by myself and created together with Anita Meraviglia and Luca Pelicci under my guidance.

4.4.1 Analysis methods

As seen in section 4.3, signals from the elastic scattering of neutrinos on electrons in liquid scintillator experiments are indistinguishable from background events, including radioactive decays occurring inside of the scintillator. While this method allows a low-threshold real-time detection of solar neutrinos, it thus also makes it impossible to separate signal from background events on an event-by-event basis. However, because of the different energies and physics behind each signal and background component, spectral shapes can vary wildly, which allows us to perform spectral fits (shape analysis) when using energy estimators. This is still complicated and not always possible with solar neutrinos because their interaction rates are extremely low and even relatively small contaminations of the LS or the materials surrounding it can cause the shape analyses to break down due to dominating background components. In JUNO, the very low interaction rates are somewhat compensated by the immense size of the experiment, which however also constitutes a technical challenge.

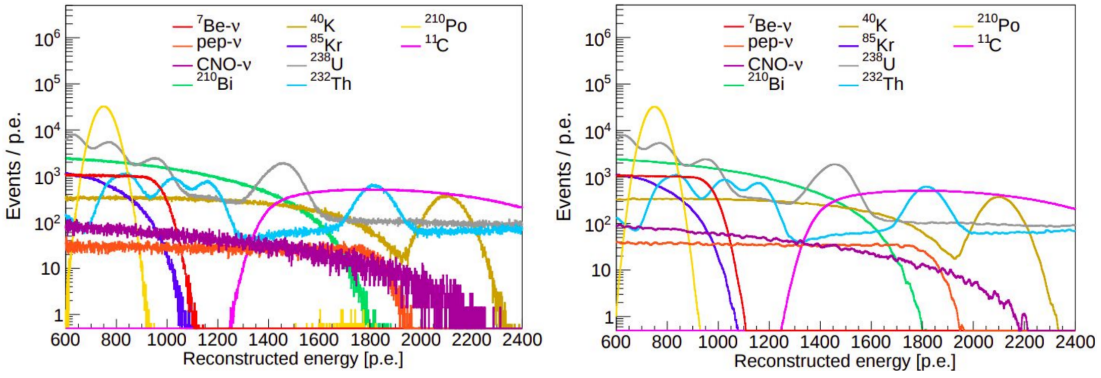
Events in this analysis are selected using the m_{NQE} variable in the range from 650 to 2400 (corresponding to about 0.45 MeV to 1.7 MeV). It is in principle the total charge (p.e.) in all PMTs in a 280 ns signal window, but with corrections. The first correction is to subtract the expected amount of dark noise per hit, and the second is meant for non-uniformity, as can be seen here:

$$m_{NQE} = \frac{\sum_i q_i/\hat{Q}_i - q_i^D}{\sum \hat{\mu}_i(R)/\epsilon_i} \cdot Y, \tag{39}$$

where the sums are over all PMTs, q_i is the measured charge on the PMT, $\hat{Q}_i$ is the expected charge for one photon hit, q_i^D is the expected dark noise, $\hat{\mu}_i(R)$ is the reconstructed-radius-dependent positron-equivalent energy expected at 1 MeV, ϵ_i is the relative detection efficiency, and Y is the light yield in p.e./MeV [240]. In terms of the JUNO software (jargon), it is an energy reconstruction used in OMILREC (QTMLE) with a RecTimeLikeAlg non-uniformity correction. The lower bound of the ROI is above the maximum energy of pp neutrinos and ^{14}C , though it includes some ^{14}C- (and ^{210}Po-) related pile-up (see section 4.3.4), and the upper bound is right where ^{10}C background starts to increase. Further, as discussed in section 4.3.2, a fiducial volume cut of 14 m is applied to all events as this is enough to ensure that no external events fall inside of the ROI.

In order to perform the aforementioned spectral fits, probability distribution functions (PDFs) were generated for each possible signal and background contribution. When scaling these PDFs to different expected rates, one obtains a dataset such as the one for the *baseline* scenario seen in fig. 32. The PDFs were saved as histogram objects with an

integral set to one. Their bin contents indicate the probability of a detected event from this type to produce a number of p.e. falling within the limits of the respective bin. The detection efficiencies, which can be determined using the Monte-Carlo simulations, also need to be taken into account since not all physical events produce a signal in the detector.



(a) (b)

FIGURE 33: Energy spectra of different components for one year of data-taking in the *baseline* scenario. (a) before and (b) after the Savitzki-Golay filter was applied, see text.

Because of the enormous quantities of data expected in JUNO, it was not possible to simulate enough events to calculate the shape of many of the PDF with sufficient precision, even for just a year of data-taking. This is a problem because statistical fluctuations can exacerbate or even create spectral structures out of the blue which can then artificially increase sensitivity. One solution to deal with this dilemma is to smooth the available PDFs. This is somewhat difficult because a line must be drawn for the strength of the smoothing. Too strong and it would erase the statistical fluctuations but also distort the data past recognition thus biasing the results, too weak and the statistical fluctuations remain problematic (also biasing the results). One must also make sure that the smoothing procedure does not change the innate shape of the underlying curve. An optimisation was thus performed by Davide Basilico[13] using a Savitzki-Golay filter [245] to maximise the reduction in high-frequency noise of the fluctuations (which can be seen easily using a FFT) while keeping the absolute difference between the original and the smoothed distribution under certain thresholds [246]. The results of this procedure can be seen in fig. 33.

Estimating JUNO's precision (*a.k.a.* sensitivity) to the different solar neutrino rates was then performed by creating many pseudo-datasets and fitting them individually. This was all done using the Jülich nUsol Sensitivity Tool (JUST) [247], see section 4.4.2. All results were also cross-checked with MUST - an equivalent software tool from a

[13] INFN Milano

group at INFN Milano. Each dataset was generated by individually sampling from all relevant PDFs and adding all components together while considering TFC assumptions. As was seen earlier, the datasets are divided in two histograms each. The reason for this is the TFC algorithm (see section 4.3.3). The number of events for each component and each pseudo-dataset was sampled from a Poisson distribution centered around the expected number of events (based on radiopurity assumptions and exposure), see tables 8 and 12. One should note that the aforementioned smoothed PDFs were used in both the generation of the datasets and as shapes when performing the spectral fits. Systematic uncertainties related to uncertainties on the shapes were therefore not taken into account, neither were other systematic effects.

In the fits performed in these estimations of the sensitivity, each signal and background component was given a single parameter even though both histograms are fitted simultaneously. This is of course possible because the ratio between the exposure of both the TFC-sub and TFC-tag histograms is known. Each parameter represents the sum of the number of corresponding events in both histograms. Because of the difference in exposure between both datasets, a conversion factor is used internally to convert the parameters from the total number of events related to the parameter to histogram-specific events. The detection efficiency and ratio of events falling into the ROI must also be taken into account in this conversion. Exceptions were made for ^{11}C, where one parameter was implemented for each histogram (the efficiency of the TFC algorithm is not a-priory known), and for ^{10}C and ^{6}He, which were only included in the TFC-tag dataset because their event rates in the TFC-sub histogram are negligible.

An example analysis result can be seen in fig. 34. The corresponding full correlation plot with all parameters can be seen in appendix C. As one can see, in this example the solar neutrino rates can are well reconstructed - with Gaussian distributions around the injected values. In this case the analysis is successful and JUNO's precision to a given solar neutrino can be determined as the standard deviation of the corresponding distribution. In some cases - for example when backgrounds are too high or overall statistics are too low, the results can however show biases, in which case the standard deviation is no longer appropriate to represent JUNO's sensitivity - which is therefore refrained from in this thesis. Biases can also arise when correlations between species are too large. This can in some cases be remedied by increasing the sizes of the datasets, but in some cases one needs to impose external constraints on at least one of the concerned parameters (for example when the required amount of statistics is unrealistic - a common problem in experimental neutrino physics). Using the standard deviation is also not appropriate in cases where the reconstructed distributions are not Gaussian. This for example occurs when one component "rails" to zero: meaning that a large portion of the

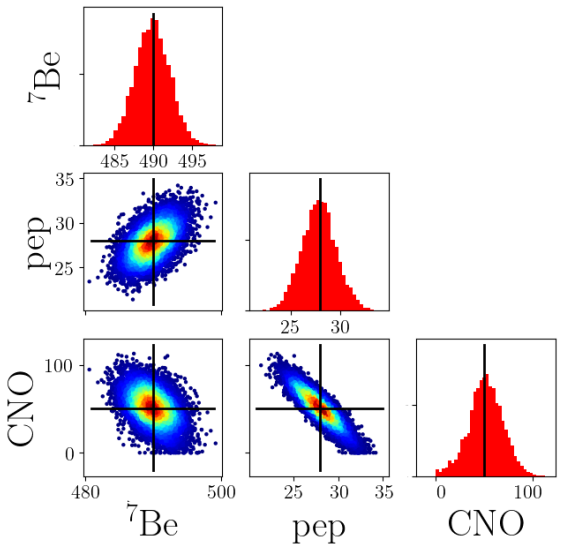


FIGURE 34: Example analysis result for solar neutrinos with 10.000 fits in the *baseline* scenario with six years of data-taking. The black lines indicated the injected values that were used in the generation of the data, and each point in the histograms and two-dimensional plots was obtained with one fit to a generated dataset (see text). All values are in cpd/kton. As one can see the values were reconstructed with no significant bias. JUNO's precision to said solar neutrino for this scenario can be extracted by taking the standard deviation of the corresponding red histogram.

time its rate is reconstructed as zero. This also indicates that in this scenario, JUNO is not sensitive to that component.

4.4.2 Fit software development

Software & architecture As previously mentioned, JUST [247] is a fit software or *sensitivity tool* that I designed for this analysis and thesis. The software was written in C++17, using the ROOT::TMinuit minimiser in order to perform the aforementioned fits. It was also developed using a GitLab server with extensive use of continuous integration tools, making it impossible to accidentally end up with a version that has unexpected effects on the results. Finally, because the software is deployed using CMake, it is very easy to install once the correct versions of ROOT and of the C++ compiler are on the target machine. While some parts of JUST were specifically designed for JUNO's solar analysis, it was in general designed to be used as a multi-purpose fitter. Controlled entirely with text-based configuration files (and of course input data) it can with be used very little effort for many different applications. In general, JUST can be used to fit an arbitrary number of data histograms with any given set of PDFs, given as input, arbitrarily distributed over the histograms. This means that a single PDF can be used in multiple histograms. Further, while in general one fit parameter corresponds to a single PDF, one can just as easily use a single parameter for different

PDFs or for the same PDF shared in different histograms. Great effort was also put into making JUST memory-safe and scalable, meaning that very large numbers of fits can be performed in a single instance (up to $\mathcal{O}(10^6)$ when using standard $\mathcal{O}$(GB) RAM) without risking overflow. Since JUST does however not make use of parallel processing, it is usually desirable to perform only up to a few thousands of fits in one instance, and to have several instances running simultaneously. The data histograms can either be read directly from ROOT files (as are the PDFs) or be generated from the PDFs under different assumptions on their event rates, exposure, TFC performance, and detection efficiency. In the latter case, JUST will automatically generate a user-defined number of Poisson-fluctuated datasets, fit them, and store the results of the fits along with some metadata. The program is fully modular and its structure is as such:

1. **Parser** starts by reading command line arguments which must specify the path to the different files. The files are then read and the relevant information is extracted. This includes the type of fit, the location of possible ROOT files, details of the fit including *e.g.* exposure, information about the different PDFs to be used in the fit and, separately, the PDFs to be used in eventual pseudo-data generation. *Parser* creates a global *NuFitConfig* object, accessible from all the other modules, which is used to store this information.
2. **DataReader** opens the different ROOT files specified in the *NuFitConfig* objects, and reads out histograms which are all converted to C++ vector objects. This is done because it was observed by Ömer Penek that this results in significant speed gains in the fitting itself. *DataReader* also gathers information about the binning of the different histograms and PDFs to make sure that they are consistent.
3. **ToyDataGenerator** uses random numbers to sample from the different PDFs, according to the exposure and rates that were specified in the config files. It is only called in case pseudo-datasets are to be used. Instead of creating all pseudo-datasets before the fits happen, *ToyDataGenerator* creates a generator with enough instructions to be able to create one pseudo-dataset. This generator can then be used by other modules in order to generate pseudo-datasets on the fly - this ensures that there is never more than one dataset in memory.
4. **Fitter** initialises a TMinuit instance and when needed calls the ToyDataGenerator in order to perform fits. The design of the *Fitter* is based on the interplay of two classes: *NuFitContainer* and *MinuitManager*. The former contains all the information needed for the fits (e.g. the data and the likelihood function to be used) and as its name indicates, the latter is used to actually perform the fits.
5. **FitResults** is used to extract the relevant information from the fit output in *Fitter*. It can combine parameters from different sets (e.g. mixing between parameters that were fixed in the fits and parameters that were left free to vary) and work

with covariance matrices to output uncertainties with or without consideration of their correlations.

6. **OutputManager** uses the output from *FitResults* to create different plots. It also opens output files, creates ROOT Trees, and saves all relevant information.

Binned template likelihood fit As previously mentioned JUST performs binned likelihood fits. Since the parameters in the fits are the integral contents of the different PDFs (the number of events associated with each PDF), it is also in practice a template fit. Since any Poisson likelihood includes factorial terms which can quickly become computationally unstable, small approximations were performed following [248]. As such the likelihood has the following form:

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{i=0}^{N_{bins}} y_i(\boldsymbol{\theta}) - n_i + n_i \cdot (ln(n_i) - ln(y_i(\boldsymbol{\theta}))), \tag{40}$$

with

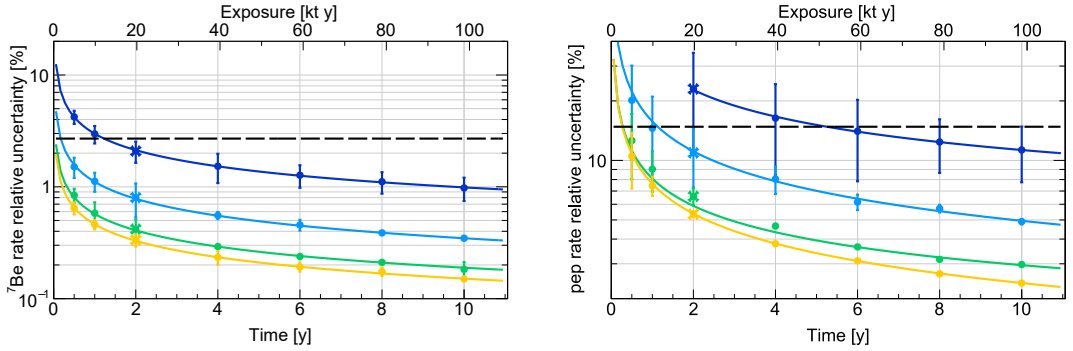
$$y_i(\boldsymbol{\theta}) = \sum_{j=0}^{N_{species}} \boldsymbol{\theta}_j \cdot PDF_{i,j}, \tag{41}$$

where $\mathcal{L}$ is the log-likelihood, y_i is the fit function estimated by the parameters $\boldsymbol{\theta}$ in the bin i, $PDF_{i,j}$ is the value of the PDF of species j in bin i, and n_i is the number of data entries in bin i. In case multiple histograms are included in the fit, the sum in eq. (40) will simply extend over the other histograms as well.

Using JUST JUST is a C++17 executable which is launched with different configuration files that can be specified as command line arguments. The first such file, *general options*, is used to give information about the input files and conditions for the fit (*e.g.* exposure, likelihood, whether to generate pseudo-datasets and how many, etc.). The second, *specieslist*, contains the name of all PDFs to be used in the fit (the name of the corresponding histogram in an input ROOT file). The fit parameters used in the fit are entirely and automatically deduced from the *specieslist* file. This is achieved with the user not only giving the name of the PDF, but also the parameter name, as well as details about the parameters to consider in the fit, such as the minimum and maximum bounds of the parameters as well as initial guesses. Should two parameters be used for the same PDF, one could give the same PDF name twice with different parameters names. The opposite is also possible. One can also use *specieslist* to specify whether a parameter should be fixed or have an additional Gaussian pull-term. Finally, *toy rates* contains information about the pseudo-dataset generation, *i.e.* which PDFs to include and with which expected rates (in cpd/kton) to generate them.

4.5 Results

All results presented in this section were obtained by performing $\gtrapprox 10^4$ fits with both JUST and MUST for many different conditions. By using both pieces of software the results could be cross-checked by both groups independently. Most comparisons in this section were done versus the so-called *standard configuration*: one year of data taking (exposure of 9.87 kt yr), tagging power and subtracted exposure of 0.9 and 0.7, respectively, and the fits were performed in a region from 650 to 2400 in the m_{NQE} variable. As explained in section 4.4.1, in each fit there is one free parameter for each species. Exceptions exist for ^{11}C which has one parameter per histogram, and for ^{6}He and ^{10}C which were not included in the TFC-subtracted spectrum. When estimating the sensitivity on CNO some constraints (Gaussian pull terms) were sometimes used, as will be explained in section 4.5.3.



(a) (b)

FIGURE 35: Expected relative uncertainty on (a) ^{7}Be and (b) *pep* neutrinos as a function of time, for different radiopurity scenarios. The errorbars show the mean and RMS of the uncertainty distributions from the fits in each point, and the lines result from fitting the errorbars with $A + B/\sqrt{\text{Time}}$ and are only for display purposes. The scenarios are described in section 4.3.5 and the black dotted line represents the best results from the Borexino experiment [46]. As one can see, JUNO is expected to surpass Borexino's results within about one year of measurement, except for *pep* neutrinos in the IBD scenario. Given JUNO's $\approx$ ten year lifetime expectancy, JUNO is expected to always reach unprecedented levels.

4.5.1 ^{7}Be neutrinos

Given the high expected rate of ^{7}Be neutrinos in JUNO (about 500 cpd/kton) and their characteristic shoulder-like spectral shape, it is relatively easy to measure. JUNO's expected sensitivity to ^{7}Be neutrinos can be seen in fig. 35(a). As one can see, after about one year of data-taking JUNO is expected to surpass Borexino's precision in all scenarios. While it is expected to barely reach the Borexino levels in a year for the *IBD*-scenario, one can see that for the *baseline* we expect to reach a level of 1%, and about

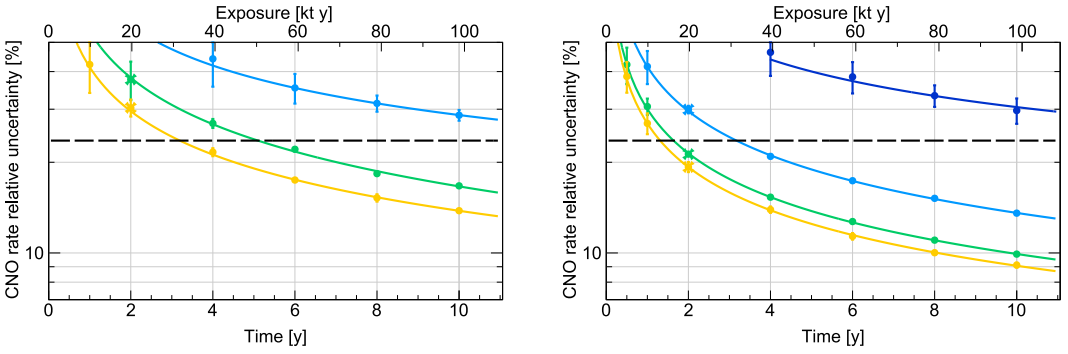
0.6% for *ideal*. After ten years the precision reaches 1%, 0.3%, and 0.2%, respectively. These results are however of course dependent on possible unforeseen backgrounds and out-of-equilibrium components inside of the scintillator. The background which could affect the ^{7}Be results the most is ^{85}Kr, which could rise to very high levels following a simple air-leak or emanation from the AV. We investigated the effects, up to about 5×10^{6} cpd/kton, of additional ^{85}Kr in the different scenarios. We found that it starts affecting the ^{7}Be uncertainty significantly around 1×10^{5} cpd/kton where it shifts it up to 2% (in the otherwise standard configuration) in the *Borexino* and *Ideal* scenarios, which would still be better than any previously existing measurement. After that the uncertainties start growing exponentially, up to about 12% for all scenarios. Another important source of background is ^{210}Po which we found is only really problematic in the *IBD* scenario. Concretely, even an additional 1×10^{6} cpd/kton contamination of ^{210}Po would not make the results less precise than those of Borexino in all other scenarios. In principle, initial ^{210}Po would not be not hugely important because of its half-life of about 4.5 months, but since Borexino also observed *supported* ^{210}Po coming from ^{210}Pb (see section 4.3.1), it is important to monitor. Further, pile-up affects JUNO's sensitivity to ^{7}Be neutrinos, particularly ^{14}C-^{210}Po pile-up. Indeed, it was found that pile-up increases the uncertainty on ^{7}Be neutrinos by about 0.5% in the standard configuration. However, the rates of ^{14}C and ^{210}Po are both measurable, be it through second-cluster events [189] or pulse-shape-discrimination, respectively. It was found that a 5% external constraint on the rate of pile-up, which is realistic but necessitates a detailed MC treatment, is enough to render the effect of pile-up negligible. Lastly, *out-of-equilibrium* ^{226}Ra, a sub-chain element of ^{238}U, could also already approximately double the uncertainty on ^{7}Be neutrinos with just a 7 cpd/kton contribution (which could occur within JUNO's design requirements). This does however not significantly change for higher ^{226}Ra contaminations, even with a factor ten more. This initial change in sensitivity can be understood as originating from an additional nuisance parameter in the fit which is highly anti-correlated with ^{238}U, a large background source. For higher rates, its characteristic alpha peaks can however very clearly be identified in higher-energy parts of the (TFC-sub) spectrum, which limits additional sensitivity loss.

4.5.2 *pep* neutrinos

The expected flux of *pep* neutrinos is about 50 times lower than its ^{7}Be counterpart, and its spectral shape is much less pronounced. Additionally, large parts of its spectrum, which is very similar (in shape and rate) to that of CNO neutrinos, are covered by ^{11}C events. Consequently, one cannot reach the same levels of sensitivity as for ^{7}Be, but one can still expect JUNO to reach unprecedented areas. The sensitivity results

for *pep* neutrinos are depicted in fig. 35(b). Here again, JUNO is expected to surpass Borexino's precision after about one year of data-taking. However, in the *IBD* scenario, one needs about two years before the analysis starts to be able to give unbiased results. In that case it will also take on average about six years to surpass Borexino's levels, though these sensitivity results have a large uncertainty themselves. In the *baseline*, *ideal*, and *borexino* scenarios, JUNO is expected to reach a precision of about 6.5%, 4%, and 3%, respectively, in six years of data-taking. The most important factor for JUNO's sensitivity to *pep* neutrinos is the achieved performance of the TFC algorithm: we observed a factor two in the *pep* uncertainty over the range of TP and SE that we investigated. The exception again is the *IBD* scenario in which the high rates of ^{238}U and ^{232}Th - which both cover much of ^{11}C's energy range - make the removal of ^{11}C events almost insignificant. For the other scenarios, we further saw that the TP parameter is in general much more important than the SE - showing that this analysis is also important for future decision-making concerning the TFC algorithm's performance.

4.5.3 *CNO* neutrinos



(a) (b)

FIGURE 36: Expected relative uncertainty on CNO as a function of time for different radiopurity scenarios, with (a) pep as a fully free parameter in the fit and (b) pep constrained (see text). The errorbars show the mean and RMS of the uncertainty distributions from the fits in each point, and the lines result from fitting the errorbars with $A+B/\sqrt{\text{Time}}$ and are only for display purposes. The *IBD* scenario is not shown in (a) because it was not possible to obtain a reliable measurement with it. In (b) however JUNO starts to obtain sensitivity, albeit a low one, even in the *IBD* scenario (shown in dark blue) after about four years. One can see that JUNO can reach unprecedented levels even without *pep* constraint if the *ideal* scenario is reached. The requirements for a sub 20% measurement are much lower in case of a *pep* constraint.

The detection of CNO neutrinos is very difficult for several reasons. The obvious one is their low expected rate compared to pp-chain neutrinos. Additionally, because their total rate ends up being similar to that of *pep* neutrinos - whose spectral shape is also very similar to that of CNO, there is a significant anti-correlation between the two

components in our fits. The same is true for the shape of ^{210}Bi, which is an internal radioactive background that could be in JUNO even for long-term measurements, see section 4.3.1. However in JUNO, the large difference between the (higher) expected rate of ^{210}Bi will probably make the correlation less significant. Finally, a large part of CNO neutrinos spectral range is covered by ^{11}C decays. Thanks to the TFC algorithm however, this effect is expected to be small compared to the aforementioned backgrounds. JUNO's expected sensitivity is shown in fig. 36(a). As one can see, unprecedented levels of 20 % can be reached after about six years in the *ideal* and *Borexino* scenarios. The sensitivity for the *IBD* scenario is not shown because even for long periods of time, it is not possible to obtain a reliable measurement in that case. This however shows that, in *ideal* conditions, the enormous statistics that can be expected from JUNO could be sufficient to allow a measurement of CNO neutrinos, even without further constraints as had to be done for a discovery in Borexino. The rate of *pep* neutrinos can however be easily constrained to a level of 1.4% by using solar luminosity measurements [249], the global fit to solar neutrino data, and the *pp* over *pep* ratio through *pep* neutrino measurements [250] (see section 3.4 for a more detailed explanation). This allows to break a lot of the degeneracy between those parameters, and the results including this constraint can be seen in fig. 36(b). As one can see, this improves the expected sensitivity down to 10% over ten years of data-taking in the *ideal* scenario. It also makes it possible to achieve unprecedented results with JUNO even in the *baseline* scenario after four years, which is extremely encouraging. With the *pep* constraint one could also measure CNO neutrinos in the *IBD* scenario after four years of data-taking - however in this scenario JUNO would not be able to reach unprecedented precision. As an outlook, because of the large amounts of statistics, JUNO might even be able to separate between the ^{15}O and the ^{13}N components of CNO neutrinos. At the time of writing this is still under investigation, with preliminary results suggesting that in the *ideal* scenario, the former could be determined with a 10% precision and the latter with about 20%, and that both could be measured with a factor 1.5 more uncertainty in the *baseline* scenario. Additionally, while it is yet unclear how this would be achievable in a large detector such as JUNO, an external constraint on the rate of ^{210}Bi, as was done for the CNO discovery in Borexino (see section 3.6), could help improve the results further.

5 OSIRIS Calibration

Like any other detector, OSIRIS (see section 2.3) will need to be rigorously calibrated in order to translate observables into physical properties. The calibration of OSIRIS will be performed with two systems: the source insertion system and the light injection system. In the context of this thesis, large efforts were put into the development of the source insertion system. This section presents the strategies and methods used for OSIRIS's calibration.

In the case of OSIRIS, calibration is multi-faceted since the detector itself is composed of 76 PMTs, all encased in a complex geometrical structure, which must first themselves be calibrated. This is done in combination with a laser and a LED system. The subsequent calibration of detector-wide variables necessitates reliance on a Monte-Carlo simulation software in which said geometry in encoded. This involves an energy calibration and a vertex reconstruction calibration, done with radioactive sources to account for non-uniformity. Further, radioactive sources are also used in OSIRIS in order to monitor the properties of the scintillator with low latency. Section 5.1 outlines the different goals of the calibration which take into account the detector design (see section 2.3.2), sections 5.2 and 5.3 describe the different calibration systems, section 5.4 delves into charge and timing calibration and the strategies behind it, section 5.5 does similarly for the energy and vertex reconstruction calibration, section 5.6 explains how a low-activity ^{40}K source will be used to monitor the scintillator, section 5.7 details the calibration campaign or how the aforementioned calibrations will be performed. Further, section 5.8 delves into the on-site software used for the calibration, and section 5.9 offers a discussion of the expected performance and critical tasks for the OSIRIS calibration.

5.1 Calibration goals

This section outlines, in the order in which they should be calibrated, the different detector properties that necessitate calibration.

PMT timing The OSIRIS detector consists of an array of PMTs. These individual detectors are the first obvious target for calibration. The chips inside of the PMTs allow us to sample their ADC every two nanoseconds, which over repeated measurements makes it possible to be sensitive to even small effects like the TTS which for the OSIRIS PMTs is of about 1.2 ns. Thus the first goal of the calibration is to align the PMTs in time with a precision that does not significantly worsen the innate resolution given by the TTS.

PMT charge Since the height of a photo-induced pulse on a PMT depends on different parameters that include some properties of the PMT setup itself (*e.g.* gain, background noise such as thermo-electron emission from the photocathode, see for a detailed example [251]), it must be calibrated as well for each PMT. This so-called *charge* calibration refers to associating a number of initial photo-electron peaks as measured in ADC channels. In practice the integral of the peak is often used as a variable for this calibration since it is proportional to the number of photo-electrons. This calibration can be done by measuring many pulses with a low intensity such that single photo-electron (s.p.e) peaks can be observed. This provides a reference for how much *charge* one photo-electron generates in each PMT.

Multi-p.e. peaks While higher-intensity measurements can also be used for charge calibration, a large s.p.e. sample is in any case necessary and also offers a great way to study the shape and behaviour of waveforms. In the case of OSIRIS, one expects an average of two-to-three p.e. per PMT for a typical measured event - it thus makes sense to have multi-p.e. samples for robustness. In addition, the time-response of PMTs is expected to be different for higher charges [192].

Vertex reconstruction Because of $\mathcal{O}(15\%)$ inhomogeneities in the energy response of OSIRIS depending on the position of events in the acrylic vessel, vertex reconstruction is an important part of the total reconstruction of event properties in OSIRIS. It is also explicitly very important since it is needed to design fiducial volume cuts, as well as to look for Bi-Po or ^{85}Kr events, see section 6.2. This is also a powerful tool to cross-check OSIRIS's Monte-Carlo simulation, see section 2.3.2, which is especially important in OSIRIS since not all points inside of the fiducial volume can be reached by the source calibration system, see section 5.2.

Energy reconstruction At this point the question remains of how many measured photo-electrons correspond to which initial particle energy. This is most typically achieved (as will be explained later) using radioactive sources with a known decay energy. Different energy estimators will be used in OSIRIS, a straightforward one being *e.g.* the sum of all p.e. for each event. The energy calibration will then assign an energy value (in MeV) to each measured p.e. value. As will be seen in section 5.2.3, OSIRIS's energy calibration will be performed with gamma sources. It is important to note that the energy scale of other particles like alphas or beta is in general different in the liquid scintillator, for example because of quenching.

Scintillator monitoring Using all aforementioned calibration results, one can also infer some properties of the liquid scintillator itself. These properties could change over time, particularly because of the constant inflow of new scintillator in OSIRIS. Part of the purpose of the calibration systems was therefore chosen to not only calibrate the detector once, but also to give the possibility of repeated measurements. For some properties, as will be elaborated further in the section 5.6, the source calibration system will even be able to provide continuous monitoring.

5.2 Automated Calibration Unit

5.2.1 Overview

The Automated Calibration Unit (ACU) [202] is a fully sealed and autonomous machine that was originally designed to be used in the Daya Bay [206] experiment near Hong Kong. The Daya Bay collaboration produced 25 ACUs and used 24, keeping one as a spare, which is the one that was provided to the JUNO[14] Collaboration for use in OSIRIS. Pictures of the ACU can bee seen in fig. 37.

The main purpose of the ACU is to lower different capsules (see section 5.2.3) directly into the AV. It is mainly composed of a stainless steel plate (*a.k.a.* turntable) with a 240° range of motion, which supports three acrylic wheels holding a capsule each. The revolver-like rotation of the turntable can place each wheel over a hole in the bottom flange of the ACU. The motors on the individual wheels then slowly lower the capsules along FEP-coated cables through an acrylic pipe (called the calibration pipe) which feeds the sources directly into the AV. Under ideal circumstances, the bell jar (see bottom left in fig. 37) would be closed once the radioactive isotopes are installed, and then kept closed until OSIRIS's end-of-life.

5.2.2 Off-axis ACU position and implications

As described in section 2.3.2, there is a diffuser near the top centre of the detector which is used to divert the flow of scintillator. For this mechanical reason, sources cannot pass through this position. In fact, the possible positions for the ACU range from a radial distance of about 90 cm to 130 cm from the detector centre. This innate asymmetry of the calibration axis presents challenges as it forbids us to really use the symmetries of the detector for calibration.

[14] A huge thank you goes to the Daya Bay collaboration, Jianglai Liu and Feiyang Zhang, and to the whole SJTU group for all the generous help!

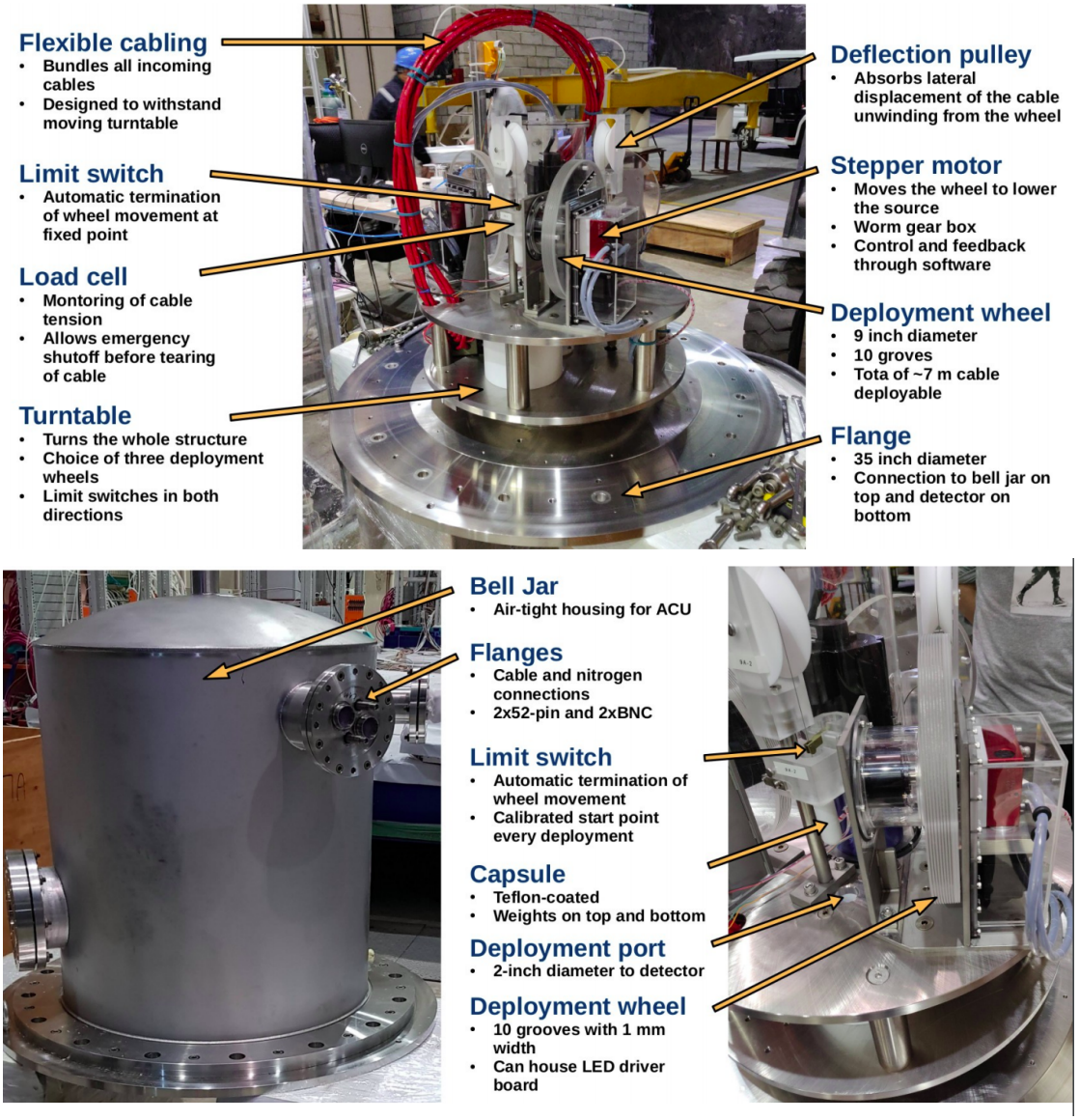


FIGURE 37: The Automated Calibration Unit used in the OSIRIS Experiment. The pictures show the ACU in its opened and closed state - with a zoom on one of the capsule/acrylic wheel setups in the bottom right. *Figure credit*: Christoph Genster

As the radial position of the ACU had to be decided relatively early on in my work on the OSIRIS calibration, the decision was made using only simulation data from radioactive sources. While in practice it could have benefited from input on the side of now-existing LED simulations, it was nevertheless found that this configuration also allows all LED-related calibration goals to be achieved, see section 5.4. These early studies were repeated with up-to-date simulations which led the same conclusion. The results can be seen in fig. 38. The idea is that since we can't place the calibration capsule in the centre of the detector, and must heavily rely on the Monte-Carlo simulation, the best we can do is to choose an axis such that the simulation can be cross-checked as thoroughly as possible. This results in an optimal position at $\rho = 120\,\text{cm}$ (where ρ is the

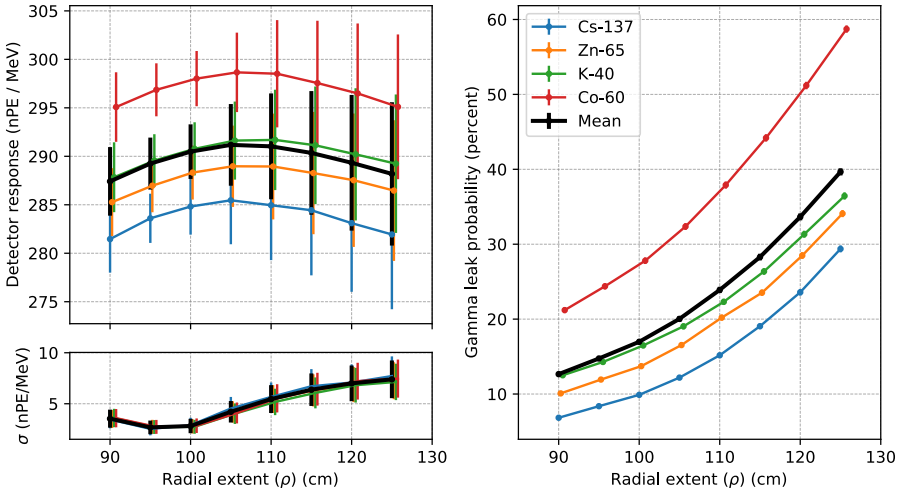


FIGURE 38: Top left: Measured detector response for different gamma sources as a function of ρ. The values are calculated as the average of regularly taken samples from -110 cm to 110 cm on the z-axis. The black points represent the average over all sources. The errorbars show the standard deviation of each distribution. Bottom left: standard deviations from the top left plot, the errorbars represent the uncertainties on the sampled standard deviation. As one can see, the detector response is different for different gamma sources - this is due to energy non-linearity. Right: *gamma leakage* fraction as a function of the radial position. *Gamma leakage* refers events that deposited at least part of their energy outside of the liquid scintillator. As one can see, the closer one is to the edge of the AV (150 cm), the higher the variation of the sampled light yield values. However, being too close to the edge implies a loss of visible light that gets exponentially higher with increasing radial distance.

radial distance from the detector centre), where the variation of all sampled detector responses is close to maximal while keeping *gamma leakage* under control. *Gamma leakage* is a loss of detectable light caused by gammas that exit the scintillator before having deposited their entire energy. This calibration axis has the added benefit of sampling from a wide range of light yield values which are representative of conditions throughout the detector. For practical reasons, the real radial position of the calibration axis in the detector turned out to be 121.5 cm.

5.2.3 Capsules

The ACU is equipped with three capsules which can be lowered into the detector one-by-one, using mechanisms and safety measures that are described in section 5.2.4. The three sources are hereinafter referred to as: 1) the multi-gamma source, 2) the potassium source, 3) the LED capsule.

The potassium source The potassium source, see fig. 39, is designed to be inside of the AV over long periods of time. The goal, as will be elaborated in section 5.6, is to have a low-activity source that can be used to monitor the properties of the liquid scintillator in near real-time while not disturbing other physics analyses. It was therefore decided to insert a $\approx 1\,\mathrm{Bq}$ source inside of the detector. ^{40}K is a perfect candidate since its low natural abundance makes the construction of a reasonably-sized capsule easy, and because its 1.46 MeV gamma decay is close to the median energy we expect to observe in OSIRIS. Indeed, one can calculate the activity of the ^{40}K capsule using:

$$A = \frac{ln(2)}{t_{1/2}} \cdot \frac{V\rho}{M} \cdot \epsilon \cdot BR, \tag{42}$$

where A is the activity in Becquerel, $t_{1/2}$ is ^{40}K's half-life (1.25 Gyr), V is the volume of the source material which simply consists of natural potassium, ρ is the density of potassium ($0.89\,\mathrm{g\,cm^{-3}}$), M is the mass of one potassium atom (39.96 Da), *epsilon* is ^{40}K's natural abundance (0.012%[15]), and BR is the branching ratio of ^{40}K's gamma decay (10.72%). When reversing this equation, one obtains:

$$V = \frac{A \cdot t_{1/2} \cdot M}{ln(2) \cdot \rho \cdot \epsilon \cdot BR}. \tag{43}$$

This results in a volume of about $320\,\mathrm{mm^3}$ for a ^{40}K activity of 1 Bq. This is reflected in the designed source volume of the capsule (see fig. 39 for more details), which is at the time of writing under construction in the SJTU[16] workshop thanks to Feiyang Zhang. A prototype was already tested in an experimental setup at the Daya Bay site.

The multi-gamma source This capsule holds three radioactive isotopes at once: ^{137}Cs ($E_\gamma = 0.66\,\mathrm{MeV}, t_{1/2} = 30\,\mathrm{yr}$), ^{65}Zn ($E_\gamma = 1.12\,\mathrm{MeV}, t_{1/2} = 244\,\mathrm{d}$), and ^{60}Co ($E_\gamma = 2.43\,\mathrm{MeV}, t_{1/2} = 5.3\,\mathrm{yr}$). The combined activity is of about 10 kHz and the relative (gamma) activities were chosen according to the relative branching ratios of the respective isotopes such that the activities of all isotopes be equal when averaging over OSIRIS's lifetime. To outline the numbers, one can look at the fraction ϵ between the average activity over a given period of time from t_1 to t_2 and the starting activity (at $t = 0$):

$$\epsilon = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} N_0 e^{-ln(2) * t/t_{1/2}} dt, \tag{44}$$

where N_0 is the number of mother nuclei at time $t_0 = 0$, and $t_{1/2}$ is the half-life of the isotope. There are of course uncertainties given the complex planning and possible unforeseen problems during the construction of any experiment. The radioactive capsules

[15] https://en.wikipedia.org/wiki/Potassium
[16] Shanghai Jiao-Tong University

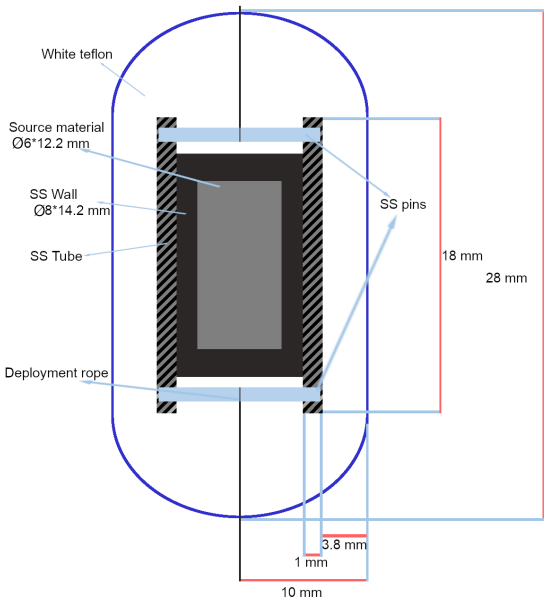


FIGURE 39: The capsule holding the potassium source. The dimensions of the capsule are such that the outer diameter does not exceed 20 mm, to make sure that it fits in the calibration pipe. Additionally the volume of the source material must be large enough to hold 1 Bq worth of ^{40}K. This takes into account the abundance of ^{40}K in natural potassium as no enrichment is needed/performed for this. SS refers to stainless steel.

were all ordered through *HTA Co. Ltd*[17] and are expected to arrive just shortly before they are needed to avoid wasting any source activity while waiting for other systems. Let's assume that we will be able to start measurements about two months after the capsules have been delivered to the JUNO on-site.

Isotope	Half-life (months)	BR (%)	E_γ (MeV)	ϵ
^{137}Cs	360.2	85	0.66	0.99
^{65}Zn	7.4	51	1.1	0.63
^{60}Co	63.2	100	2.5	0.95

TABLE 13: Isotope properties for the multi-gamma source. ϵ was calculated using eq. (44) with t_1 = two months and t_2 = eight months. BR refers to the branching ratio of the gamma decay and E_γ to the gamma's energy.

Using the numbers from Table 13 and setting the average activities to be equal, one can infer that the starting activity of the sources should be 2.75 kBq, 2.9 kBq, and 4.35 kBq for ^{137}Cs, ^{60}Co, and ^{65}Zn, respectively. This results in an average activity of 2.75 kBq over the six-month filling of JUNO, with final activities of 2.75 kBq, 2.66 kBq, and 2.05 kBq at the end of the filling period, respectively. The calculations were slightly simplified by assuming the activity of ^{137}Cs to be constant over time. It is to be noted that these activities only refer to the gamma decays. Other possible beta or alpha decays are of no importance since the isotopes will be delivered in Teflon capsules that fully

[17] Thanks to our colleagues at IHEP for facilitating the ordering process

shield the outside world from these particles. The only exception is for ^{65}Zn, where β^+ decays occur which can produce two 511 keV annihilation gammas. In practice though, the branching ratio is quite small and, see section 5.5, one can usually separate these gammas from the main gamma decay branch in subsequent analysis. The design of the capsule itself is very similar to that of the potassium capsule, see fig. 39, but with a cylindrical source volume of 4 mm in diameter and 4 mm in height instead.

The ^{60}Co isotope decays into an excited isotope of Nickel which usually emits two γ rays with a total energy of 2.5 MeV. Theoretically, if one of the isotopes looses energy by interacting in the Teflon capsule or outside of the scintillator volume, a second peak around 1.1 MeV to 1.3 MeV would be observed. While this could in principle interfere with the ^{65}Zn peak at 1.12 MeV, it was observed in simulation studies that it does not in practice significantly affect the performance of the calibration.

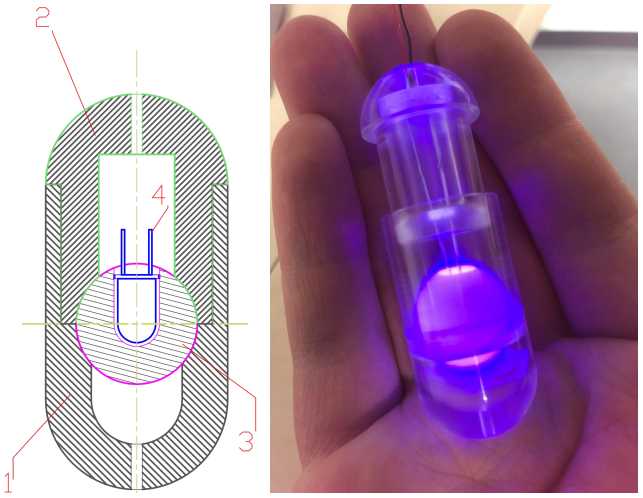
The LED capsule is an acrylic capsule used to carry a LED inside of the LS volume where it will be pulsed for timing and charge calibration of the PMTs. It is produced at the IKP workshop in Jülich but it is based on an original design by the Daya Bay experiment [202]. It consists of two acrylic parts that are glued together to hold a diffuser in which the LED is inserted. The diffuser is made of Teflon, with a special surface treatment to maximise isotropy[18]. SciGrip Weld-On 4SC glue is used to glue the acrylic parts together as it provides a strong grip, can fit in tight spots due to its low viscosity, and stays transparent after bonding.

The LED inside of the capsule is powered with a driver shown in fig. 41. It takes a negative DC power that is used to control the height of the pulses, and triggers on a TTL input. The driver itself resides in an acrylic wheel of the ACU and is controlled from a distance.

5.2.4 Mechanisms and safety measures

All ACU materials were selected for low radioactivity and chemical compatibility with the LS: Polymethyl methacrylate (PMMA), Polytetrafluoroethylene (PTFE), and stainless steel 316. The acrylic wheels are each equipped with grooves to keep the cables from moving horizontally, which could add uncertainty to the position determination. Each wheel can hold up to seven meter long cables - FEP coated stainless steel cables (McMaster Carr 3423T22) for the radioactive sources and to hold the capsule weights, and a FEP-coated mini-coax cable (Cooner Wire CW2014-3650 F) for the LED.

[18] Big thanks to Raphael Stock and his colleagues in the Technische Universität München workshop!



(a) (b)

FIGURE 40: The LED capsule used in OSIRIS. (a) shows a drawing with 1- an upper acrylic part, 2- a lower acrylic part, 3- a Teflon diffuser, and 4- the LED itself. This drawing is not to scale. (b) shows a picture of a LED inside of a diffuser inside of a capsule. The capsule is not closed because of the long "legs" of the LED, but of course in the real setup, those legs will be cut such that the capsule can close. One can see that the blue LED light is emitted from all directions from the diffuser, instead of just in the forward direction.

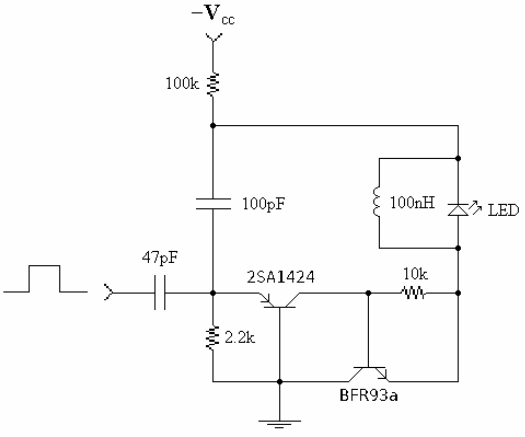


FIGURE 41: A drawing of the driver used to pulse the LED, based on [252, 253]. The driver takes a negative DC voltage as constant input, used to charge the capacitors and effectively control the power of the LED pulses. The other input on the left is a TTL whose rising flank is basically used to trigger a LED pulse by switching two transistors on, which quickly discharges said capacitors, over an inductance, into the LED. This was shown (see section 5.4) to produce pulses with a width (sigma) of less than 2.24 ns.

The four axes (turntable + three wheels) are each equipped with a stepper motor, an optical encoder, and a worm gear box. This allows for a very precise control over the positions of all sources, with an uncertainty of less than 1 cm on their (z-)position. The ACU is outfitted with several redundant safety measures, to ensure that nothing can be dropped inside of the detector and that in general no need arises to re-open the ACU, thereby risking to compromise the radio-purity of the liquid scintillator. Firstly, the worm gear boxes cannot be back-driven, which ensures that the sources don't drop down (due to gravity) in case of a power failure. Secondly, all axes are equipped with limit switches (one for each wheel motor, two for the turntable) which automatically terminate any movement of the corresponding motor when triggered. Further, load cells on the three source axes ensure the constant monitoring of the tension on the cables. Our control software (see section 5.8 for more details) automatically stops motor movements and raises alarms in case of any discrepancy. Two weights, about 100 g each, are mounted on top and below each source thereby using gravity to keep the cables straightened at all times. The former is necessary to counteract the buoyancy of the sources in the liquid scintillator. The maximum speed of the motors is set relatively low to avoid any sudden jumps in cable tension. Additionally, while the lowest breaking strength of the cables was measured to be of at least 2 kg [202], the torque of the stepper motors is limited to less than one kg strength-equivalent. To ensure that the cables stay in their respective grooves on the calibration wheels, extra Teflon rollers, which limit the radial movement of the cables with respect to the acrylic wheels, are placed next to the wheels. Finally, a CCD camera is mounted on the top of the bell jar, inside of the ACU, pointing at the source injection hole.

The ACU is connected to a control PC and a nitrogen gas interface via two double-sided ISO feed-through flanges. The former is equipped with two half-inch nitrogen connections for in- and out-flow, respectively. The latter holds two 52-pin electrical cables and two BNC connections. The electrical cables are connected to all the different electronic components inside of the ACU *e.g.* the stepper motors and load cells. The first BNC connection carries the CCD camera image, and the second is used to carry TTL pulses to pulse the LED driver.

5.3 Laser system

The laser system for OSIRIS was designed and tested by Tobias Sterr from the Eberhard Karls Universität Tübingen. A description of the system is included here for completeness. The Laser system can calibrate the iPMTs on a sub nanosecond level, and use the s.p.e. data it so gathered to also perform a charge calibration. The idea is to use a

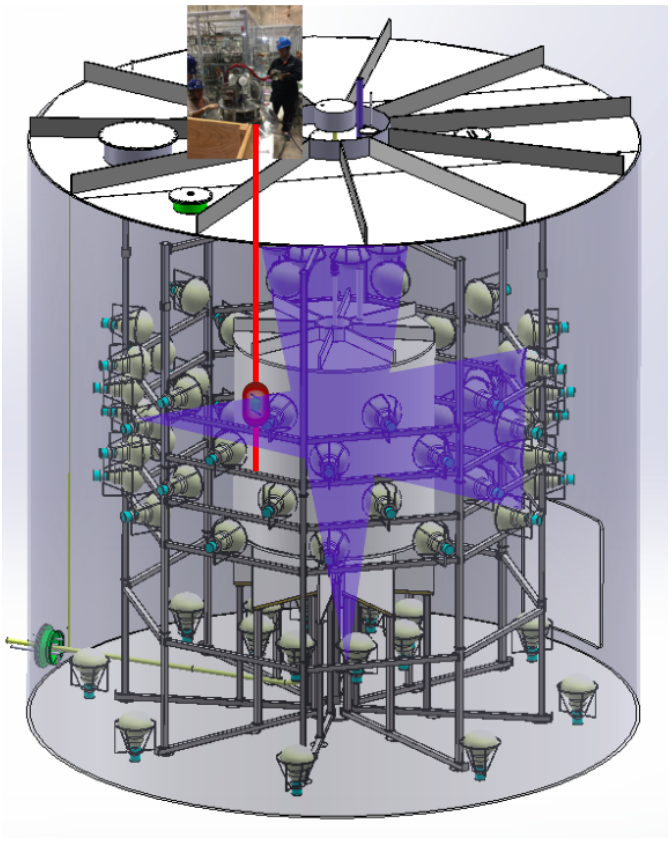


FIGURE 42: Simplified illustration of the calibration systems. The ACU can be seen at the top left, injecting source capsules into the AV through the calibration pipe (superimposed by a red line representing the cables). The blue cones stem from two different possible laser diffuser positions that are situated at different points on the PMT frame.

pico-second laser (PiL042X from ALS GmbH) to generate pulses that are then routed through different optical fibres to illuminate the iPMTs. A graphical illustration of the idea can be seen, alongside the source system, in fig. 42.

To elaborate, the laser signal is passed through a cascade of optical switches and fibre splitters. During calibration, a total of 24 different fibres can then alternatively be chosen to send pulsed laser signals through. Each fibre leads to a specially designed diffuser placed on the PMT frame. A drawing of a diffuser can be seen in fig. 43. The positions of the diffusers were optimised, using the simulation software, to guarantee a similar light level on each PMT. The length of all fibres must thereby be as equal as possible to avoid differences in light travel time between them, at least to a negligible level compared to the TTS. It was calculated that 2000 s.p.e. events are needed per PMT in order to achieve the sought accuracy of 25 ps in the inter-PMT time offsets. Because of the low light-intensity required to obtain mostly s.p.e. peaks, this results in the system needing to generate about $2 * 10^5$ pulses per diffuser. Given the frequency

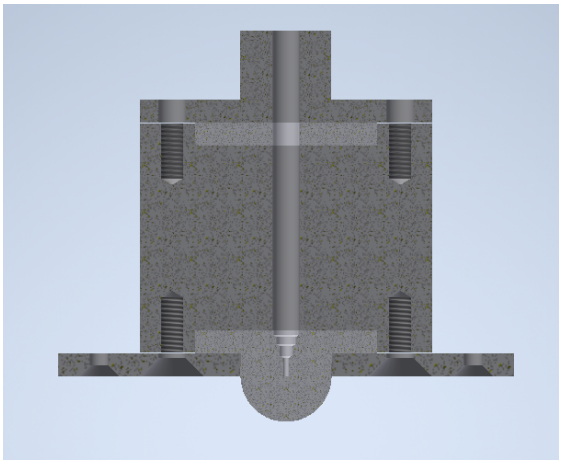


FIGURE 43: Cross section of a laser diffuser. The fibre can be seen in the centre, other materials are PTFE (light grey) and stainless steel 316L (dark grey) to comply with OSIRIS's strict radiopurity guidelines. *Source:* [207].

with which the laser can be pulsed, this results in the possibility of calibrating all PMTs in less than 10 minutes, allowing the laser calibration to be performed once per day.

5.4 PMT Charge and timing calibration

This section describes the strategies behind the timing and charge calibration of OSIRIS using the source calibration system. Since its requirements are much stricter it focusses on the timing calibration aspect. Section 5.4.1 quantitatively explains the goals of the calibration, section 5.4.2 explains how timing information is extracted from the data and goes into the simulation work performed in this analysis, then section 5.4.3 shows how the calibration plan was calculated while satisfying all requirements, and finally section 5.4.4 shows the results and concludes.

5.4.1 Goals

This section outlines the goals that should be achieved and boundaries that must be respected during the timing calibration. Quantitative requirements on the timing calibration can be seen below:

1. Maximum occupancy below 0.01. As mentioned in section 5.2.2, it is not possible to illuminate all PMTs with the same intensity. However, in order to get a consistent time response from the PMTs, it is necessary to keep them at the same light level (see fig. 18 in [192]). The solution is that the occupancy, defined as the average number of photons measured per LED pulse by a PMT, must be much lower than one - it then follows from the Poisson distribution that the number of s.p.e. events

will outnumber the two-or-more p.e. ones. An occupancy μ of 0.01 corresponds to a 0.5% chance of getting a two-or-more p.e. event relative to getting a s.p.e. one. Since the requirement is on the maximum occupancy, there will in general be one PMT with exactly $\mu = 0.01$, and all other PMTs will have an even lower one. The problem that can arise is that many more LED pulses will be needed until enough data can be gathered to get significant results for the furthest PMT. One must thus find a balance.

2. Uncertainty of 0.1 ns on the reconstructed time offsets. The time offsets are the object of the timing calibration, whose goal is to minimise the time differences between the PMTs. This requirement was chosen empirically.

3. Low bias on the reconstructed time offsets. Any possible biases introduced by the timing calibration should be small compared to the TTS (about 1.2 ns).

4. Since the reconstructed time offsets depend on the position on the LED due to reflections in the AV (see section 5.4.2), each PMT should be calibrated with data from several points along the calibration axis, for robustness. In this analysis this number of positions per PMT is set to four.

5. The so-acquired data should be usable for a precise charge calibration.

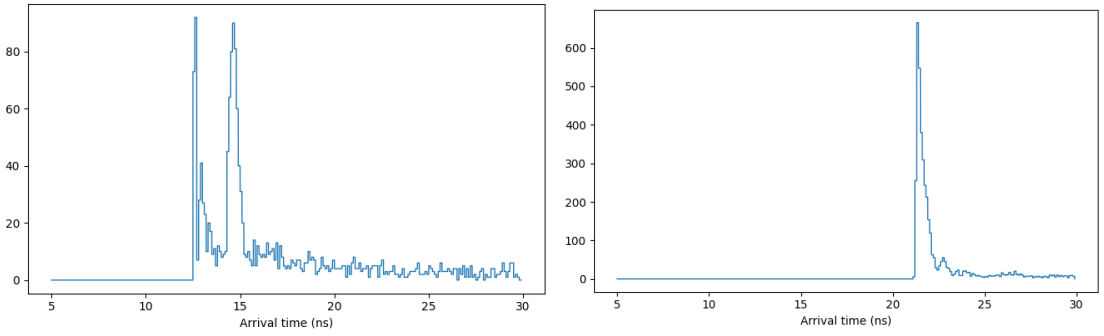
5.4.2 Methods

Simulation data As part of this analysis, a special simulation mode, specifically for the LED calibration, was introduced inside of OSIRIS's geant4 Monte-Carlo software, see section 2.3.2. In this mode, the LED capsule is built according to the latest technical drawings, as are the cables lowering it inside of the detector and the weights used to counteract buoyancy from the capsule. The simulation emits light from the LED with an angular distribution which follows measurements that were taken by Tobias Sterr[19], see Appendix D.1. This is not trivial because these measurements were not performed in liquid scintillator but in air. The measured angular distribution therefore includes a PMMA→air transition which is not present inside of the detector. In order to counteract this, the LED capsule in the simulation was implemented as a homogeneous capsule made of air (standard geant4 composition) with a thin layer of acrylic around it. The photons are then generated inside of a volume, corresponding in position and shape to the volume of the teflon diffuser, which lies at the centre of the air-capsule. The photons are randomly oriented based on the measured distribution, travel through air (without interacting with it), then transition to PMMA, thereby practically reversing their PMMA→air transition, and finally enter the scintillator, thus simulating the

[19] Eberhard Karls Universität Tübingen

PMMA→LS transition. This is expected to provide a good approximation of the true angular distribution. In the simulation the light then propagates inside of the LS, using a realistic propagation model in a realistic detector, see again section 2.3.2, and for each simulated event it saves all photon hits on PMTs. It is to be noted that the the simulation does not include OSIRIS's electronics but merely outputs a list of photon hits with their respective timestamps, for each PMT and event. A large amount of simulations were launched in the context of this analysis. Specifically, data was simulated for 25 different capsule positions with 500.000 pulses every 10 cm from −120 cm to 120 cm. The simulations were repeated for different amounts of primary photons per LED pulse (constant in each simulation run), with $n_{prim} = [1, 10, 50, 100]$.

Reflections Before reaching the PMTs, photons emitted from the LED must travel through the scintillator and traverse the AV. Because of the optical properties of these two materials, this can lead to total reflection. This was indeed observed to occur frequently for light coming from the LED, see the example in fig. 44. The figure shows that while for some LED positions the light to a given PMT only takes a straightforward path to the PMTs, for some others many of the photons will have been reflected at least once. This results in a double-peak structure in the photon hit-time distribution.



(a) (b)

FIGURE 44: Example of PMT first-hit times (non-smeared MC-truth information) for the same PMT, with the capsule at (a) z=−100 cm and (b) z=100 cm. In the simulation, the photons are generated in the LED capsule at the time t=0. As one can see, in (a) there are two peaks very close to each other, the latter being understood to originate from reflections on the acrylic vessel. In (b) one can see that the light had to travel for a longer time, but that although a second peak can also be seen, it is too small to significantly affect any timing analysis.

Getting timestamps When two hits occur simultaneously on two different PMTs, one will in general measure different arrival times. This is of course in part due to the TTS, but also due to the electronics of the respective PMTs. On average, each hit will thus be artificially moved by a PMT-dependent value. The goal of this calibration is to measure this value between all PMTs to allow us to correct for it. This can be

done by measuring many LED-induced hits - and extracting one timestamp from the hit distributions on each PMT. To re-iterate this important point: the goal is to obtain the average PMT-induced time shift, not an individual value for each photon. In order to get realistic timing information from the simulation, each output file (one for each position and n_{prim}, see above) was processed as described in the following. First multi-p.e. hits were removed, as in real data one would also only use s.p.e. data for timing calibration. Given the lack of an existing waveform reconstruction in OSRIS, an existing perfect waveform reconstruction was assumed - a good enough approximation given reasonable expectations. Secondly, a smearing according to assumptions on the real experiment was applied, as summarised in table 14. Finally the time-of-flight of each hit was subtracted.

Source	LED electronics	ACU	TOF	OSIRIS
Value	measurement	0.5 ns	0.4 ns	1 ns

TABLE 14: Smearing added to simulation timestamps. The "LED electronics" values are randomly sampled from a time distribution that was measured using an OSIRIS PMT in a setup in Aachen, see Appendix D.2, and already includes all LED-related electronics as well as PMT TTS. "ACU" was measured from the ACU electronics, "TOF" stands for time-of-flight and was conservatively estimated using the MC, and "OSIRIS" is a conservative estimation of the noise induced by other OSIRIS electronics. The values indicated in nanoseconds indicate that random values are sampled from Gaussians with said standard deviation.

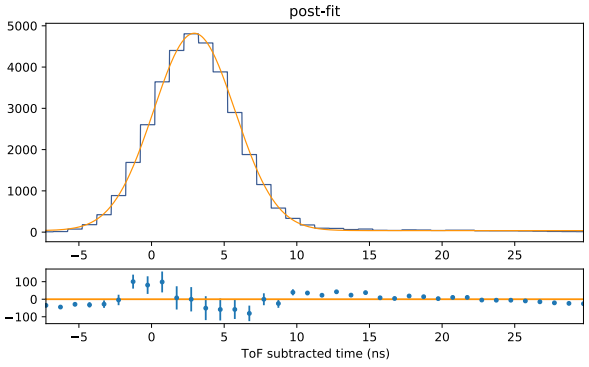


FIGURE 45: Example fit to photon arrival times on a PMT. The data is from a LED capsule at $z = 0$ on the calibration axis and from a single PMT. This summarises the arrival times of all first-hits recorded on the PMT from 500.000 LED pulses with 100 photons each. The fit function used in this plot is the sum of a Gaussian and a constant. While the residuals show systematic effects, the fit procedure was found to yield consistent results over many samples.

After the aforementioned smearing and time-of-flight correction, the hits were put in a histogram and fitted with the sum of a Gaussian and a constant. The latter was introduced because it empirically improved the results. The peak position of the Gaussian is in the following used as the reconstructed time for each PMT. An example fit can be seen in fig. 45. This also illustrates a problem: the fine structures caused by reflections

are to small to be seen on realistic (smeared) data. In order to see the effect of the reflection on the reconstructed timestamps, one can define:

$$T = t_R + t_{offset}, \tag{45}$$

where T is the extracted timestamp from the fits, t_R is the bias induced by reflections, and t_{offset} is the offset from PMT electronics which we wish to calibrate. For each simulation file, I (blindly!) added a random offset sampled uniformly from -10 ns to 10 ns to each PMT. The so-obtained data was then fitted as previously described for each position and PMT. After subtracting the random offset again after the fit one thus obtains t_R, and the results can be seen in fig. 46. As expected t_R is very position dependent, clearly not negligible, and must therefore be carefully considered in the following analysis.

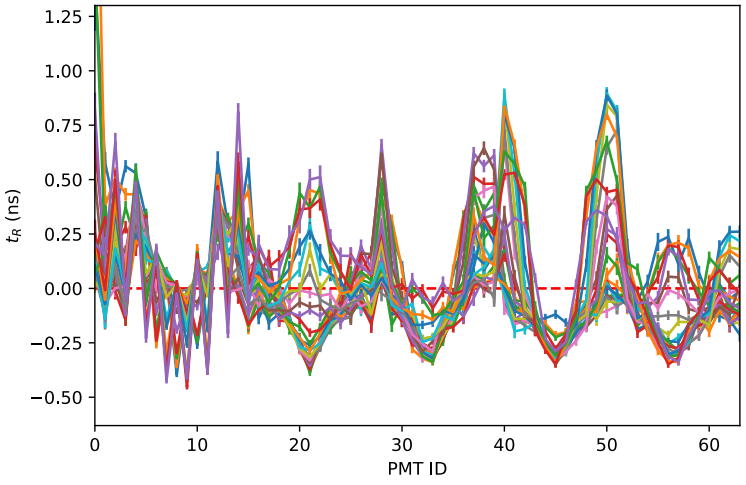
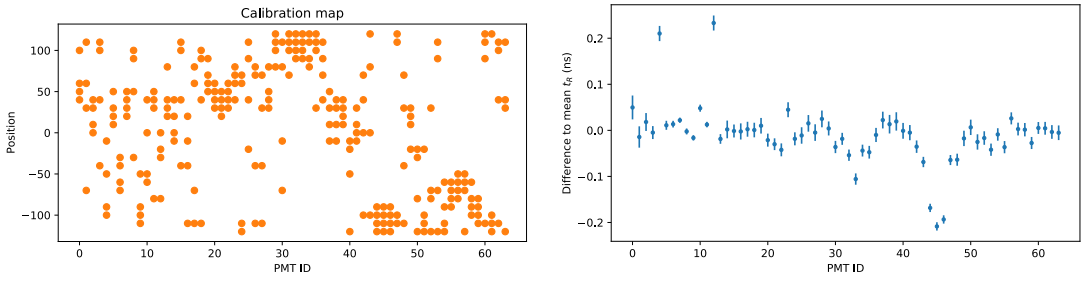


FIGURE 46: t_R (see text) as a function of the PMT ID. The different colours correspond to different positions of the LED. As one can see, t_R is very position-dependent. This can be well understood when looking at fig. 44. Because of the high amount of reflections on the inner side of the AV, light can often take several routes and still reach the same PMT. This will then obviously depend on the position - and this figure shows to what extent this affects the timing calibration.

Calibration map This section presents a method to minimise the biases caused by reflections. The main idea is that by combining data performed by the LED at different positions, it is possible to find PMT-position pairs such that the overall effect of reflections is minimised. Since the timing calibration is only interested in minimising the differences of the time offsets between the PMTs, the absolute value of t_R is not important, instead the relative differences between t_R values must be minimised. In order to choose these positions, a procedure was developed in which we consider that the calibration of each PMT should include data from several LED positions. This increases

the robustness of the calibration and also allows on-the-fly cross-checks of the simulation expectations - which will help to eventually adjust it once on-site and increase our trust in our understanding of the detector. As such, the procedure starts by calculating the mean value of t_R for all PMTs and positions: μ_{t_R}. For each PMT, the N_{pos} positions resulting in the smallest value of $|t_R^i - \mu_{t_R}|$ are then chosen, whereby t_R^i is the value of t_R for a given position (and for that PMT). For each PMT, the mean $\bar{t}_R$ over all chosen positions is then used as the reconstructed time. The results can be seen in fig. 47. As one can see, the procedure yields a t_R distribution where the maximum bias between PMTs is below 0.4 ns. In principle, one could use the same procedure and change N_{pos}. Increasing N_{pos} results in more bias as more t_R values, which are not necessarily as close to μ_{t_R}, will be used in the calculation of the mean. However, lowering N_{pos} also increases the uncertainty on said mean.



(a) (b)

FIGURE 47: (a) The so-called calibration map. The points indicate what positions are used to calibrate which PMT. The resulting t_R values as a function of the PMT IDs can be seen in (b). as one can see, this calibration map allows to keep the maximum differences in t_R below 0.4 ns.

5.4.3 Calibration plan

Number of primaries per position Given a calibration map, one can now determine the required number of primaries per pulse n_{prim}^i to be emitted at each LED position. The calculation itself is quite straightforward. For each position, one takes a look at the occupancy on all PMTs that use that position, then n_{prim} is adjusted such that the maximum occupancy is exactly 0.01. The occupancy μ on every PMT was calculated using raw simulation files (reminder: for a given PMT, the occupancy is the average number of photons recorded per LED pulse). In order to get as much information as possible from the available data, the occupancy was determined as a function of the number of primaries in each pulse. A simple linear regression was performed as this is what one would expect, and the data was found to be very compatible with this expectation. As such, for each of the aforementioned output files, one obtains two fit parameters along with their uncertainty matrix (a and b from $\mu^i = a \cdot n_{prim}^i + b$).

Time per position The calculation of the time needed at every position takes a somewhat more complicated approach. The idea is that since the uncertainty on t_R scales with $1/\sqrt{t}$, one can adjust t to reach a given uncertainty σ_{lim} on $\bar{t}_R$. In order to calculate this uncertainty as a function of the measurement time t, one must first take a look at the uncertainty as a function of the number of emitted photons. Since the uncertainties on t_R should in theory only depend on the total number of emitted photons (high pulse frequency and low n_{prim} would give the same uncertainty as low pulse frequency and high n_{prim}, given that the s.p.e. condition is kept), one can assert:

$$\sigma^i_{t_R} = \sqrt{A_i + B_i/N_i}, \tag{46}$$

where $\sigma^i_{t_R}$ is the uncertainty on the t_R value determined from a single PMT/position (see fig. 46) and N_i is the total number of emitted photons in a given measurement series. A_i and B_i are determined at each PMT/position by fitting the simulated $sigma_{t_R}$ for given n_{prim}. For a given PMT, a mean value can then be calculated over all t^i_R values from the different positions used to calibrate it. The uncertainty on this mean value can be expressed as:

$$\begin{aligned}\sigma_{\bar{t}_R} &= \frac{1}{N_{pos}}\sqrt{\sum_{i=0}^{N_{pos}}(\sigma^i_{t_R})^2} \\ &= \frac{1}{N_{pos}}\sqrt{\sum_{i=0}^{N_{pos}} A_i + \frac{B_i}{N_i}},\end{aligned} \tag{47}$$

where $\sigma_{\bar{t}_R}$ is the uncertainty on $\bar{t}_R$: the mean t_r over the positions used for that PMT and N_{pos} is the number of positions. One can now insert $N_i = f \cdot t_i \cdot n^i_{prim}$, where f is the frequency at which the LED is pulsed (1 kHz), t_i is the measurement time at the position i, and n^i_{prim} is the number of primaries emitted from the pulses at position i. To simplify the equations and because one does not expect a huge advantage from considering it, let the measurement time t_i be the same for all positions: $t_i \to T$. This results in:

$$\begin{aligned}(N_{pos}\sigma_{\bar{t}_R})^2 &= \sum_i A_i + \frac{1}{f \cdot T}\sum_i \frac{B_i}{nP_i} \\ \leftrightarrow T &= \frac{1}{f}\frac{\sum_i \frac{B_i}{nP_i}}{(N_{pos}\sigma_{\bar{t}_R})^2 - \sum_i A_i}.\end{aligned} \tag{48}$$

5.4.4 Conclusion

As seen in section 5.4.3, given the requirements formulated in section 5.4.1, one can obtain a list of positions at which to position the LED. For each position, a time t_i was calculated indicating how long the LED must be pulsed at said position with a given frequency, as well as a number of primaries n^i_{prim} corresponding to the intensity with which the LED must be pulsed. These values can be seen in Table 15. As one can see, one can reach the goals of the calibration, using a LED frequency of 1 kHz, a maximum occupancy of 0.01, and a σ_{lim} requirement of 0.1 ns, within about one and a half hour. Adding the time needed for the ACU to lower and then pull the capsule back out, one gets a total of about two hours of weekly calibration time. One should note that these requirements are quite stringent. For one, it should be possible to pulse the LED up to 3 kHz, which would proportionally lower the required measurement time. Further, relaxing the σ_{lim} requirement to 0.2 ns, which seems reasonable, would reduce the required measurement time by a factor four. This calibration is therefore very possible despite the high amount of light reflection inside of the AV. When using the parameters described in Table 15, one measures between about 4800 to 35700 photons per PMT. This is a factor of 2.4 to 18 more than what is measured in the laser calibration. This s.p.e. sample can thus be used for charge calibration with correspondingly higher accuracy, which for the laser system is already 7×10^{-3} p.e.

z (cm)	t	n_{prim}
-120	2.4min±3.1s	19.0±0.11
-110	2.8min±5.2s	7.2±0.04
-100	2.8min±1.9s	6.9±0.04
-90	2.8min±1.9s	6.6±0.04
-80	2.2min±3.6s	9.6±0.06
-70	4.4min±18.2s	10.0±0.06
-60	2.2min±3.6s	9.8±0.06
-50	2.8min±1.9s	6.0±0.04
-40	2.8min±5.2s	10.0±0.06
-30	2.1min±2.4s	5.7±0.04
-20	2.4min±6.1s	38.3±0.29
-10	3.7min±7.3s	31.8±0.22
0	3.8min±5.9s	11.8±0.07
10	4.5min±6.2s	16.8±0.10
20	4.5min±6.2s	6.2±0.04
30	4.5min±6.2s	6.0±0.04
40	6.4min±20.4s	5.8±0.04
50	6.4min±20.4s	5.7±0.04
60	6.4min±20.4s	40.9±0.32
70	2.8min±5.2s	23.3±0.15
80	3.7min±4.1s	24.9±0.16
90	2.6min±1.1s	9.3±0.05
100	6.4min±20.4s	9.8±0.06
110	4.4min±18.2s	26.3±0.17
120	1.5min±1.9s	22.5±0.14
Total time:	91min±52.1s	

TABLE 15: Plan for OSIRIS's LED calibration. This plan gives for each position the requirement intensity-equivalent of the LED as well as the necessary amount of measurement time to reach the required precision (in this case 0.1 ns for each PMT.

5.5 Energy and vertex reconstructions calibration

The energy calibration of the OSIRIS experiment is performed, as mentioned previously, using a radioactive capsule containing three isotopes (^{137}Cs, ^{65}Zn, and ^{60}Co) with energies ranging from 0.66 MeV to 2.46 MeV. The ACU (see section 5.2) can deploy capsules to any position along a z-axis which is at a distance of 121.5 cm from the central axis of the detector (y=x=0). The sources will be deployed at a given position over a certain amount of time, and for each approached position the cumulative hit information of the iPMTs can be used to translate from the observable to the true deposited energy (or inversely). The energy variable used in this analysis is the number of detected photo-electrons: n.p.e. *i.e.* the total measured charge by all iPMTs for a given event. Assuming a perfect charge reconstruction, fig. 48 shows a typical spectrum as measured in a calibration run. The same position-dependent gamma events that will be measured with this capsule will be used to tune vertex reconstruction algorithms.

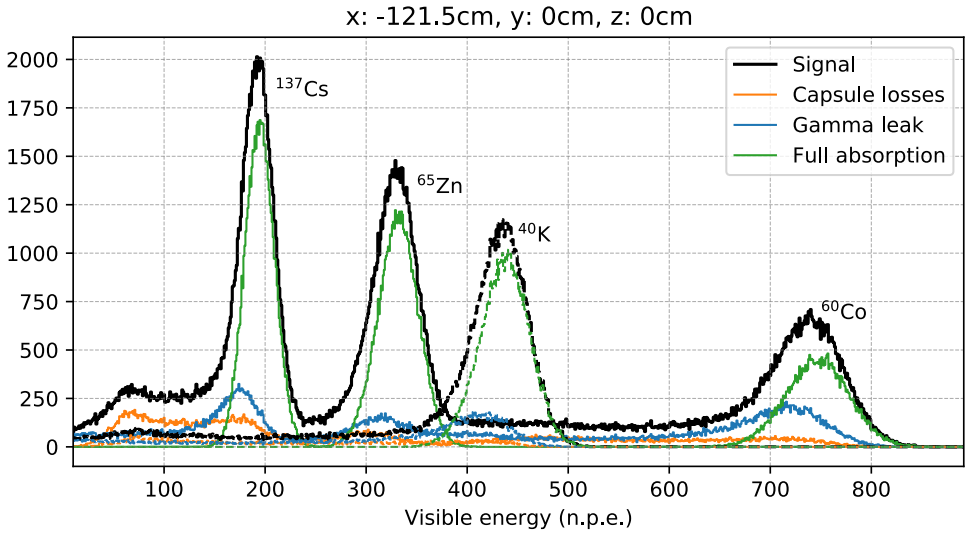


FIGURE 48: A typical calibration spectrum, simulated at z=0 cm. The solid lines show the spectrum from the multi-gamma capsule, while the dashed lines come from the potassium source. The black lines show the total measured nPE, which is what we expect to measure in real data. The coloured lines show different components. Green curves show the full absorption peak: this is what we are in principle interested in, only gammas that have deposited their entire energy in the liquid scintillator contribute here. Blue lines show *gamma leakage*: event that deposited at least part of their energy outside of the liquid scintillator. This contribution is highly position dependent. Finally, red lines show events that deposited at least part of their energy in the capsule (mostly through Compton effect) surrounding the sources. This component is position independent. The sum of the coloured bins do not equal the black curve because some events, which are not shown, belong to both *capsule losses* and *gamma leak*.

One can clearly recognise the four peaks from the three different isotopes. Care was taken in the choice of the isotopes (from JUNO's available pool [156]) to have clearly separate peaks on this spectrum. One can also see that the spectra are dominated by

three components: *full absorption*, *gamma leakage*, and *capsule losses*. *Full absorption* is what one would naively expect: the gamma deposits all of its energy in the liquid scintillator, mostly through Compton scattering, the energy gets converted to light, and the resulting peak is close to Gaussian. *Gamma leakage* refers to events in which at least part of the energy of the initial gamma is deposited outside of the liquid scintillator volume. This missing energy is not converted to light, which results in the observed low-energy tails. Finally, *capsule losses* is the result of (almost exclusively) Compton interactions of the initial gamma inside of the capsule surrounding the source material. Overall, while the *capsule losses* are identical everywhere in the detector, *gamma leakage* is of course very position-dependent in its shape and size.

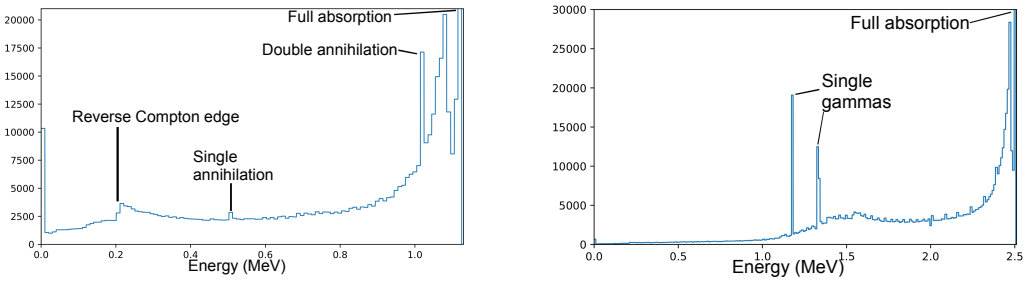
All in all, these effects result in measured peaks that are systematically shifted towards the lower energies. After all, the known energy of the isotope decays is only directly visible in the case of *full absorption*. The easiest method of energy calibration is to simply fit a Gaussian to the visible peaks. However, this will result in a systematic shift of the energy scale towards lower energies. In principle, one could apply a MC-inferred shift to every peak and to every position. For a typical calibration run (see section 5.7), since those effects are position-dependent, this would imply relying on 3*25=75 MC-inferred constants.

5.5.1 Convolution fits

To solve the aforementioned issues, I developed a convolution fit method based on [254]. The idea is that an understanding of the physics behind the energy deposition in liquid scintillator should allow us to fit the measured energy peaks in order to more reliably extract the position of the full absorption peaks. This doesn't only automatically reduce the bias we would otherwise have, but also decreases the need for reliance on the MC. It is to be noted that since the only way to test/prepare the fit procedure is on simulated data, there is of course still a dependence on the simulation, albeit to a lesser extent.

To understand the model behind the convolution fits, one can look at fig. 49. As can be seen, when looking at the energy region above the reverse Compton edge, the energy deposition can be broadly modelled as the sum of a delta peak, a constant, and an exponential increase up to the delta peak, although some isotope-dependent details remain to be included. The events in the delta peak lost all of their energy inside of the scintillator. As seen in section 1.6.5, this happens through many Compton scatterings of gammas on scintillator electrons, up until the energy of the resulting gammas reaches below 20 keV - where the cross-section of the photo-effect becomes dominant. The

[20] http://nucleardata.nuclear.lu.se/toi/nuclide.asp?iZA=300065



(a) (b)

FIGURE 49: Total energy deposited in the liquid scintillator from (a) ^{65}Zn and (b) ^{60}Co, both from the centre of the calibration axis. In both cases, the plots were cut-off on the y-axis such that the last bin (right) is not fully visible. This was done so that one can see the structure of the bins to the left of it. The last bin, which was cut-off, contains the full-absorption events. All other bins contain events which either lost energy in the capsule or outside the LS volume. The reverse Compton edge is in both plots very visible (from capsule losses), as is the exponential increase towards higher energies (from gamma leakage) - although a gap can be seen between the tip of this exponential increase and the full-absorption bin (see text). In (a) one can see an additional contribution, near 1 MeV, from two annihilation gammas stemming from a positron emission that can occur in 1.4% of ^{65}Zn decays[20][255]. In (b) one can see two peaks between 1 MeV and 1.5 MeV because ^{60}Co emits two back-to-back gammas when decaying (whose total energy is that of the full-absorption peak, 2.5 MeV).

exponential increase can be explained by the fact that the primary gamma sometimes escapes the scintillator. Because the mean free path of gammas in OSIRIS's LS is shorter than the mean distance needed to leave the scintillator, the probability of energy being carried out of the detector has the characteristic exponential decrease of gamma energy losses in passage through matter. Since the measured events are composed of individual energy depositions inside of the scintillator, the reverse is seen in the spectrum: an exponential increase of events with energy, up to the full-absorption peak (corresponding to zero leakage). Here a special feature is visible (fig. 49), namely that there exists a small ($\approx$ 20 keV) gap between the full absorption peak and the aforementioned exponential feature. This is related to the dominance of the photo-effect at this energy and below - which drastically reduces the mean-free-path and thus the probability of escape, resulting in the gap in the spectrum [254].

Finally, the constant component can be seen as an approximation of the *capsule loss* events. These are events which underwent one Compton scattering inside of the capsule surrounding the source. The visible energy in the scintillator is then the energy of the so-emitted gamma, resulting in the reverse Compton edge in the spectrum. Although this spectrum is in general not at all constant, in this case for energy regions between the reverse Compton edge and the full-absorption peak, it is a good enough approximation, particularly because the number of events that undergo this scattering is small compared to the total number of events.

Going into further isotope-related detail: in the case of ^{65}Zn an additional delta function near 1.02 MeV is needed in the model. This is the result of an unlikely decay branch of ^{65}Zn in which a positron is emitted. This positron annihilates in the capsule surrounding the radioactive material, thereby creating two gammas with 511 keV. When looking closely at fig. 49, one can even see a small contribution from single annihilation gammas near 511 keV, though the contribution is so small that it need not be included in the model. The model (delta+constant+exponential) is only valid for the higher energy part of the spectra *i.e.* above about 0.5 MeV for ^{65}Zn and 1.7 MeV for ^{60}Co.

The remainder of this section describes the mathematical formulation of the general model used to fit a peak spectrum, with the three aforementioned terms:

$$f = (a\delta|^{\mu} + bf_{const}|_{\mu_1}^{\mu_2} + (1-a-b)f_{exp}|_0^{\mu-\epsilon}) \circledast G(\mu, \sqrt{\Gamma\mu}), \tag{49}$$

where a, b are normalisation factors, μ is the position of the delta peak (full absorption), μ_1 and μ_2 are the integral limits for the constant term (capsule losses), f_{exp} is the exponential term (gamma leakage), with the superscripts denoting integral limits in the convolution, ϵ is used to model the aforementioned ≈ 20 keV gap between the exponential term and the full absorption peak, and $G(\mu, \sqrt{\Gamma\mu})$ is a Gaussian distribution, meant to model the detector resolution. The Gaussian term is centered at μ with a width of $\sqrt{\Gamma\mu}$. The term Γ is included as a "catch-all" term to model non-linearity and non-homogeneity effects. This is sufficient for our approximate model and required level of precision. Writing the convolution out:

$$f(x) = a\mathcal{F}_{gauss}(x) + b\mathcal{F}_{const}(x) + (1-a-b)\mathcal{F}_{exp}(x), \tag{50}$$

$$\begin{aligned} f(x) = a\frac{1}{\sqrt{2\pi\Gamma\mu}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sqrt{\Gamma\mu}}\right)^2} + b\int_{\mu_1}^{\mu_2} C\frac{1}{\sqrt{2\pi\Gamma x'}}e^{-\frac{1}{2}\left(\frac{x-x'}{\sqrt{\Gamma x'}}\right)}dx' \\ + (1-a-b)\int_0^{\mu-\epsilon}\lambda\frac{1}{\sqrt{2\pi\Gamma x'}}e^{\lambda x'}e^{-\frac{1}{2}\left(\frac{x-x'}{\sqrt{\Gamma x'}}\right)}dx', \end{aligned} \tag{51}$$

where C is the constant and λ is the shape factor of the exponential term. After integration the constant term becomes:

$$\begin{aligned} \mathcal{F}_{const}(x) = A\cdot C\cdot\Bigg(erf\left(\frac{\mu_2-x}{\sqrt{\Gamma 2\mu_2}}\right) - erf\left(\frac{\mu_1-x}{\sqrt{\Gamma 2\mu_1}}\right) \\ + e^{2x}\cdot\left(erf\left(\frac{\mu_2+x}{\sqrt{\Gamma 2\mu_2}}\right) - erf\left(\frac{\mu_1+x}{\sqrt{\Gamma 2\mu_1}}\right)\right)\Bigg) \end{aligned} \tag{52}$$

[20] https://physics.nist.gov/PhysRefData/Xcom/html/xcom1.html

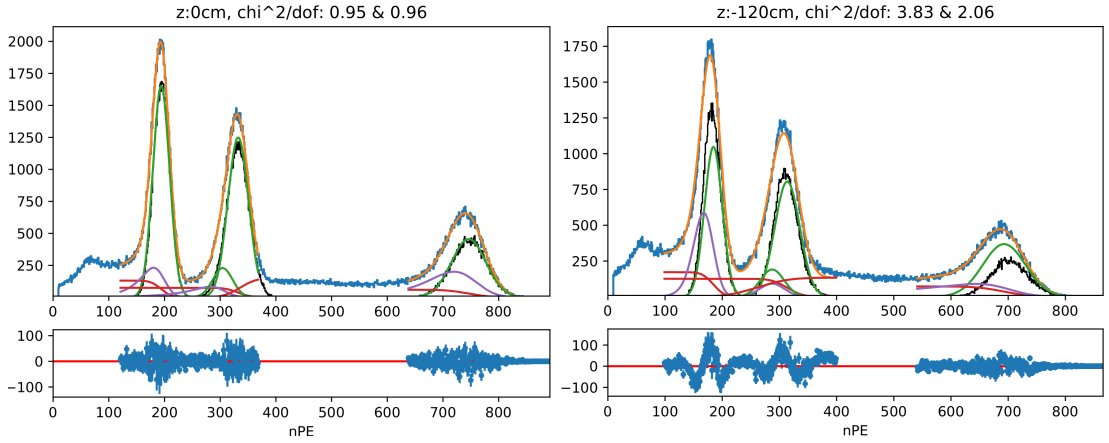
where A is a normalisation factor that can be obtained by setting the integral of the function, *e.g.* from 0 to infinity to unity. Continuing with the exponential term, after integration one gets:

$$\mathcal{F}_{exp}(x) = A \cdot \frac{1}{2\sqrt{1-2\lambda}} \left(e^{x-|x|\sqrt{1-2\lambda}} \cdot \left(erf \frac{(\mu-\epsilon)\sqrt{1-2\lambda}-|x|}{\sqrt{\Gamma 2(\mu-\epsilon)}} + 1 \right) + e^{x+|x|\sqrt{1-2\lambda}} \cdot \left(erf \frac{(\mu-\epsilon)\sqrt{1-2\lambda}+|x|}{\sqrt{\Gamma 2(\mu-\epsilon)}} - 1 \right) \right) \tag{53}$$

5.5.2 Fit example

When inserting eqs. (52) and (53) into eq. (50), one obtains the total fit function. In practice while fitting, the variables α and β were introduced, with $\alpha = a/(1-a-b)$ and $\beta = b/(1-a-b)$. The fitter was then given the variables α and β, to be converted to the normalisation terms in eq. (50) using: $a = \alpha/(1+\alpha+\beta)$, $b = \beta/(1+\alpha+\beta)$, and $1-a-b = 1/(1+\alpha+\beta)$. This allows to use unbounded variables which result in factors of a and b between 0 and 1, with the added benefit of having $1-a-b > 0$ by construction.

In the end, one needs five free parameters to fit one peak: α, β for the relative height of each component, μ for the position of the full absorption peak, then λ for the shape of the exponential term and Γ for the energy correction. When fitting the entire histogram, one additional parameter is added for the extra ^{65}Zn positron-induced peak, and two for an additional $\mathcal{F}_{const}$ term used to model the low-energy tail of the ^{60}Co distribution under the ^{65}Zn peak. This results in a total of 18 free parameters (three peaks with five parameters each + three parameters). An example fit can be seen in fig. 50. As one can see the function is a very good fit to the data. The gap between the low and high energy regions was introduced because this is a difficult region to model (it includes the reverse Compton edge of full ^{60}Co events, as well as the peaks from the individual ^{60}Co γs), while providing negligible information about the peak positions.



(a) (b)

FIGURE 50: Examples of convolution fits performed with the multi-gamma capsule at z=0 cm and z=−120 cm on the calibration axis, respectively. The blue points represent the (MC) data representing 10^5 decays per isotope, and the orange lines show the fit function. The green, red, and purple lines show the individual fit components meant to model different physical effects: *full absorption*, *capsule loss*, and *gamma leak*, respectively. Finally, the black points show the true full absorption peak as extracted from the MC. It is to be noted that the fit has at no time information about this peak, it is simply shown to demonstrate here the ability of the deconvolution. As one can see, the fits are much cleaner near z=0 cm, which is expected because it is the point with the least effects coming from the proximity to the detector edge, which is not part of the relatively simple model. However, in both cases one can see that the positions of the true full absorption peaks are reconstructed better than with a Gaussian fit.

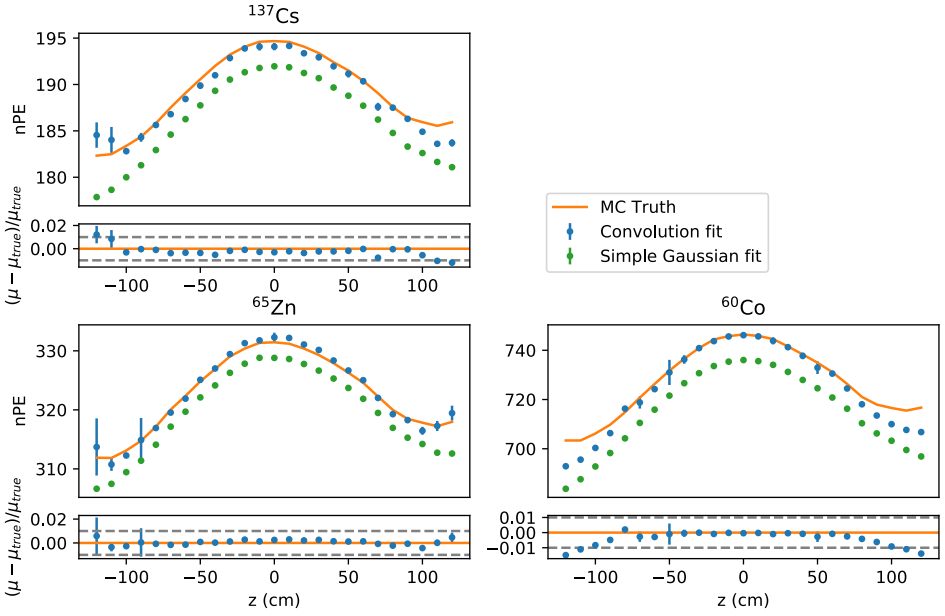


FIGURE 51: In each plot, full absorption peak position in n.p.e. as a function of the position of the capsule on the calibration axis, with relative residuals. The orange curve shows the MC truth information, the blue points show the results of the fit procedure described in this section, and the green points show the results of simply fitting Gaussians to the observed peaks. The results of the fit procedure show no systematics, except an underestimation of the ^{60}Co and ^{137}Cs peak positions near the detector edges, although in general the reconstructed values are nearly always within 1% of the true value.

5.5.3 Fit results

The fit shown in the previous paragraph was then performed on simulated data for 25 different capsule positions homogeneously distributed from $z = -120$ cm to 120 cm on the calibration axis. In each position, the energy ranges of the fits were chosen to optimise the total χ^2 of the fits (minimising $|\chi^2 - 1|$). Additionally, the values of ϵ were varied within one order of magnitude to minimise the difference between the reconstructed full absorption peaks and the MC truth information. The results of this procedure can be seen in fig. 51. As one can see, while the uncertainties from the fit procedure are higher than when simply fitting Gaussians, it basically eliminates all bias that would otherwise exist. The results of the fit procedure show no systematics except a slight underestimation of the ^{60}Co peak position near the detector edges, although in general the reconstructed value is nearly always within 1% of the true value. Other systematic effects can have an influence on the reconstruction of the energy scale as well. They were investigated, using MC samples at z=0 cm, and their results are reported in Table 16. As one can see, the only statistically significant effect is from source shadowing, but even here the bias induced in the energy scale reconstruction is of less than half a percent,

which is negligible. Source shadowing occurs when some of the light that would otherwise reach PMTs is blocked by the calibration source itself. This could be investigated in the MC by comparing data from a realistic setup to that of a setup where the source just floats in the LS. The effect of the analysis background, mostly ^{14}C and external gammas, was also investigated. The expected event rate of $\approx 40\,\mathrm{s}^{-1}$ is however very low compared to the several kBq expected from the calibration sources. Finally, the effect of pile-up was investigated. Pile-up refers to the mis-reconstruction of events that occurs when they are so close in time that they cannot be differentiated in the detector.

Effect on peak pos.	^{137}Cs [%]	^{65}Zn [%]	^{60}Co [%]
Shadowing	+0.575±0.028	+0.267±0.079	+0.436±0.032
Ana. bkg.	(-1.1±2.5)·10^{-2}	(-0.5±7.6)·10^{-2}	(-0.008±3.3)·10^{-2}
Pile-up	(-0.4±2.4)·10^{-2}	(-0.5±7.7)·10^{-2}	(+0.5±3.4)·10^{-2}

TABLE 16: Influence of different effects on the reconstructed energy. The numbers are in percent and show the relative difference between the reconstructed peak position when considering the corresponding effect or not. *Shadowing* refers to source shadowing: the capsule, weights, and cables around the radioactive source block out some of the reconstructed light. *Ana. Bkg.* (analysis background) refers to the background from the non-calibration related events in the detector, as taken from fig. 55. Most of those events are from ^{14}C and external gammas. The numbers here were calculated without taking any fiducial volume cut into account. *Pile-up* refers to experimental pile-up which occurs when two or more events happen so close in time that they are reconstructed as one event. I estimated a worst case scenario where such events can never be distinguished in the 240 ns measurement windows, and the effect on the energy scale is still negligible.

5.6 Scintillator monitoring

The goal of the potassium capsule (see section 5.2.3) is to monitor some properties of the liquid scintillator during normal operations. Any changes in the properties of the detector and particularly of the liquid scintillator would then be observable by looking at the position and shape of the potassium peak. Further, since OSIRIS is expected to be running during the commissioning of JUNO's liquid handling system - it is clear that quick feedback about the scintillator would come in extremely handy. This would not only be very useful to the people working on the scintillator, but would also inform us about the OSIRIS detector as a whole. The idea behind the potassium capsule is that it is very easy to get ^{40}K in very low concentration/activity (with a natural activity of 0.012%[21]). With a sufficiently low activity it can as such stay inside of the liquid scintillator during normal operations without disturbing the rest of the analysis. In OSIRIS, the potassium capsule will contain ^{40}K with an activity of about 1 Bq. A prototype of the capsule was tested in a Daya Bay [206] setup - showing that it does indeed contain the desired amount of potassium, with negligible contributions from other

[21] https://www.ciaaw.org/potassium.htm

radioactive isotopes. This section describes the analysis and prospects of the potassium capsule.

This analysis is about monitoring change and necessarily starts with a reference measurement of the potassium peak - to be compared with subsequent ones. The ^{40}K inside of the potassium capsule decays and produces gamma rays with an energy of 1.46 MeV, producing a peak that can be reasonably approximated by a Gaussian (because we are only interested in relative change this approximation valid in this case). The comparison between a new sample and a reference measurement is done by calculating the χ^2-difference between the newly obtained samples and a Gaussian obtained by fitting the reference sample. This χ^2 value combines position and shape information to be sensitive to any change in the potassium peak. In the following, the expected sensitivity to such changes is explored.

Data analysis In order to study the sensitivity, potassium data was simulated using the Monte-Carlo framework. For technical reasons, effects from capsule interactions were ignored, but the full potassium decay was simulated. This is relevant because a position cut is important to reject most of the background in OSIRIS, and because the gamma from the ^{40}K decay has a $\mathcal{O}(10\,\mathrm{cm})$ mean free path in scintillator. In the following, the "true position" of a given ^{40}K event is its average position - weighed by energy - of all energy depositions in the scintillator. Background was also simulated as decays from all expected internal, external, and cosmogenic backgrounds (see section 6.1 for a more detailed discussion of the background in the general OSIRIS analysis). Because of a software bug, the hereinafter referred to as "true position" of the background events is just the average position of all their energy depositions in the scintillator, not weighted. This was kept as-is in this analysis because re-simulating everything constitutes a massive task outside of the computational scope of this analysis, and because it is not expected to be as significant as to change anything fundamental.

All data points, signal and background, were then *smeared* with assumptions about event reconstruction. It would be preferable to use the actual event reconstruction methods that will be used in OSIRIS, but those don't exist yet. The assumption on the reconstruction, as was used in [208], is a component-wise (x, y, z) uncertainty of 14 cm at 1 MeV. This is scaled from the precision reached in Borexino [192] and corresponds to a 3D vertex uncertainty of about 24 cm. After the smearing, one obtains a "realistic" data sample. In principle, all events that are in a radius d_{cut} around the capsule could be selected. However, given the long measurement times needed to measure a significant amount of ^{40}K decays, one must take the flow speed of the scintillator into account. In OSIRIS, scintillator flows evenly from top to bottom with a speed of about $3.47\,\mathrm{cm\,s^{-1}}$.

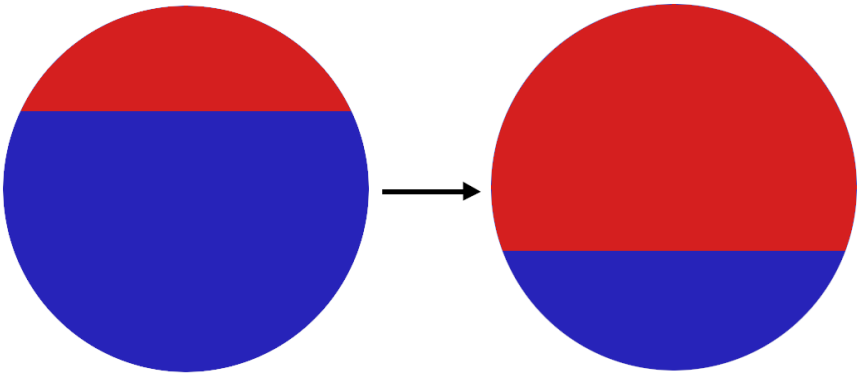


FIGURE 52: Simplified graphical representation of the cuts described in eq. (54). The red area represents scintillator tagged as "new" and the blue as "old". The arrow represents the passage of time.

The exact value will be measured very precisely with flow sensors. The flow pushes "old" scintillator down, evenly along the z-axis, to be replaced with "new" scintillator. Thus, assuming that new scintillator starts flowing inside of the sphere of radius d_{cut} when the measurement starts, then at first one would mostly select events coming from the "old" scintillator - which would cover any change up - as can be seen in a sketch in fig. 52. A time-dependent z-cut is therefore introduced in order to select only for "new" events given the hypothesis of an entry time t_0. Combined with a selection around the known ^{40}K energy region, these cuts are able to discard almost all of the hits stemming from non-^{40}K events - with an expected signal-to-noise ratio of above 700 in typical analysis ranges. Since we can safely neglect this source of background, the only background that remains is "old" scintillator being identified as "new", due to imperfect vertex reconstruction. The cuts can be formulated as:

$$\begin{cases} z > z_{capsule} + d_{cut} - v_{LS} * (t - t_0) & , t < \frac{2d_{cut}}{v_{LS}} - t_0 \\ z > z_{capsule} - d_{cut} & , else \end{cases},$$
$$\sqrt{(x - x_{capsule})^2 + (y - y_{capsule})^2 + (z - z_{capsule})^2} < d_{cut}, \tag{54}$$

where $xyz_{capsule}$ denotes the position of the potassium capsule, v_{LS} is the speed of the scintillator, d_{cut} is a cut value on the distance to the capsule, and t_0 is a variable meant to represent the time of entry of the new scintillator.

Results and discussion Using the data acquired as described above and following the Ansatz with the χ^2 calculation, one can obtain, for example, fig. 53. The orange distribution in that figure is the χ^2/dof distribution one would expect to measure if scintillator does not change. The blue distribution is what is expected if the light yield is increased by 1%. An optimisation was performed to get the cut values used in these

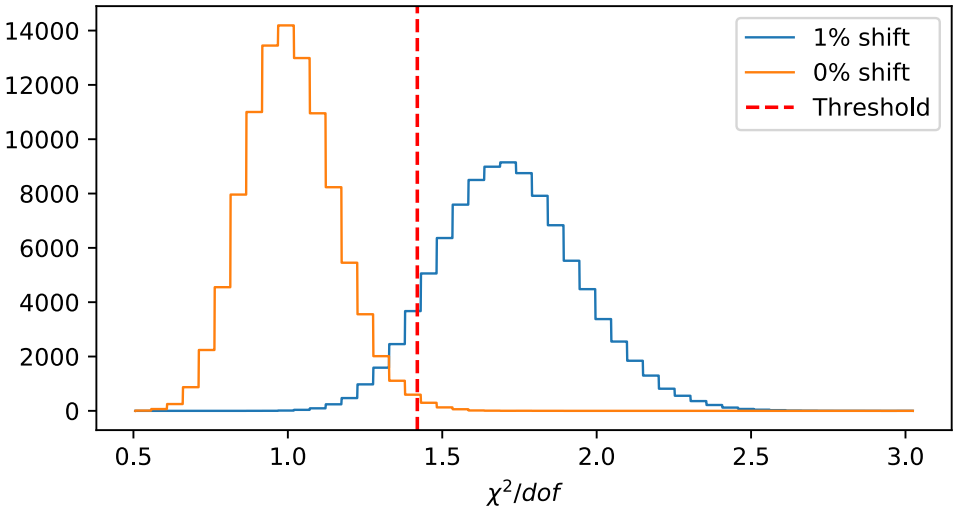


FIGURE 53: χ^2/dof distribution (see text) for a case in which the scintillator stays the same as the reference, and for a case where light yield increases by 1%. The measurement window is 4.3 h long and starts 30 min after the new scintillator has entered the $d_{cut} = 26$ cm radius sphere. Both histograms contains 10.000 toy-samples. The red line shows a value which would reject shifted data (false negative) in 10% of the cases, while only producing about one false positive a week (assuming consecutive 4.3 h long measurement windows).

results in which target type-I and type-II errors were firstly defined (false negative 10% of the time, false positive once a week), followed by a minimisation of the total time needed to achieve these results. This resulted in $d_{cut} = 26$ cm, $t_1 = 30$ min, and $\Delta T = 4.3$ h, meaning that this measurement would be obtained by waiting for 30 min before measuring for 4.3 h, assuming that the "new" scintillator enters the sphere of radius d_{cut} at time $t_0 = 0$. With this method, acceptable type-I and type-II errors, the definition of which is expected to change during OSIRIS's commissioning, can be set - along with how different the "new" scintillator is - to obtain cut values leading to the fastest possible results. This particular result gives a quantitative idea of the sensitivity of the potassium analysis. It so demonstrates that the potassium method is very powerful and can give relatively fast warnings about even small changes in the scintillator.

It should however be noted that during the actual running of OSIRIS, a single high-χ^2 value from the potassium analysis won't have consequences on its own. Rather, it is meant as a warning that something in the data has changed - and that this change should be looked at and understood before any decisions are made. An example of how much the data is affected by the conditions used in this analysis, see fig. 54. As one can see, if the high χ^2 value is from a (even only 1%) shift on the light-yield, one can expect to see it quite clearly. Further, one should keep in mind that the entry time of the "new" scintillator is in general unknown, and that the t_0 defined in the cuts rather represents a hypothesis. On the real data, several cuts with different assumptions on t_0

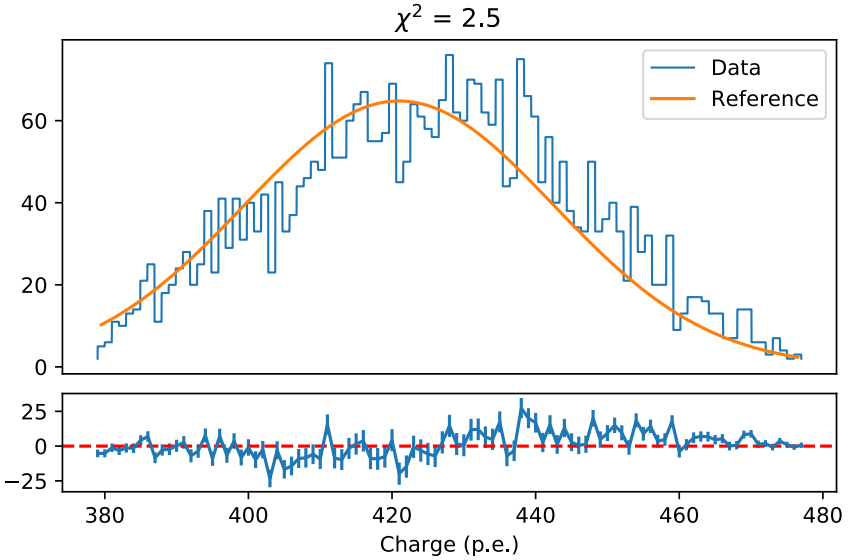


FIGURE 54: Example effect of a 1% shift in effective light yield on the χ^2 difference between measured (shifted) data and (non-shifted) reference. As one can see in this case, even a 1% shift is quite visible even by eye, particularly in the residuals. This data sample was taken under the same cut conditions described in the text (4.3 h of data-taking).

could be running in parallel to have measures with different sensitivity, which could *e.g.* allow for cross-checks. The goal is never to trace back when a change happened, only that it did.

5.7 Calibration campaign

There are three kinds of calibration campaigns planned for OSIRIS. The first, so-called "comprehensive" calibration, aims to be used only once in the commissioning phase of the experiment to provide a large quantity of data. This data can be used as a reference for future, less detailed calibrations, and to study the detector idiosyncrasies and simulation, particularly during commissioning. The second, so-called "weekly" calibration, is meant to be used, as the name indicates, on a weekly basis to monitor possible changes in the detector/PMTs. A weekly calibration can also be triggered by an alarm coming from the potassium capsule. The last kind of calibration is a daily calibration that will be performed by the laser system, which is possible since it needs less than ten minutes to calibrate all PMTs. A summary of the calibration campaigns can be seen in Table 17. These values were obtained assuming a combined activity of 10 kBq from the multi-gamma source, and a frequency of 1 kHz on the LED. The number of positions for the LED and multi-gamma sources was chosen to be 25 because the resulting 10 cm distance between points is on a similar scale than that of expected $\mathcal{O}(1\%)$ changes in the detector response. The distance between points for the multi-gamma source was

halved in the comprehensive calibration to deliver a more precise sample. This will also provide invaluable input to tune vertex reconstruction algorithms. The multi-gamma calibration is expected to generate 3.4 GB of data in the comprehensive run, and about 700 MB each weekly run. The LED calibration is expected to generate 9.6 GB per run, but since each PMT is only calibrated by the data from few points, it can in theory be reduced to about 850 MB per run. In the comprehensive calibration, the ^{40}K scan in expected to generate 3.3 GB of data.

	Source	Points	Source insertion & removal [min]	Source deployment [min]	Total [min]
Comprehensive	Laser	24	N/A	10	10
	Multi-γ	49	31	24.5	55.5
	LED	25	31	92	123
	^{40}K	9	31	540	571
	Total	107	93	10.9 h	12.5 h
Weekly	Multi-γ	25	31	2.5	33
	LED	25	31	92	123
	Total	31	62	94.5	2.6 h
Daily	Laser	24	N/A	10	10

TABLE 17: Overview of the calibration plans. The top four rows show the plan of the comprehensive calibration done at the start of operations, which will last around 36 hours. The next three rows show the plan for a weekly calibration run which takes less than three hours. These values were obtained assuming a combined activity of 10 kBq from the multi-gamma source, a potassium activity of 1 Bq, and a LED pulse frequency of 1 kHz.

5.8 On-site calibration software and DAQ

The operation of the OSIRIS Experiment will follow using an online software framework referred to as the run control. This software can communicate with all OSIRIS subsystems using EPICS[22] and can run online analyses to display live data using the so-called RootSorter[256], which is a ROOT[23] based software that was originally developed at IKP-2 in the Forschungszentrum Jülich by V. Hejny, M. Hartmann, and A. Mussgiller for the ANKE experiment. The RootSorter will be able to read data in real-time and display live plots with analysis-relevant variables. A specific window exists for the source calibration, which will also allow control over the ACU through a special UI. This UI will allow us to communicate with the run-control, which in turn will use EPICS variables to communicate with the ACU software.

The ACU software is a LabView (2015 32-bit) software that was written by Christoph Genster and myself. It is based on existing Daya Bay software but was severely refactored

[22] https://epics-controls.org/
[23] https://root.cern/

from its original state. Its main logic consists of three simultaneously running state machines. The first, the script loop, awaits instructions (e.g. from the run control) then converts them to LabView-readable commands to cycle them to the command loop, which is the second state machine, when it detects that it is ready. The command loop takes specific motor instructions and runs them while making sure to handle any possible problems. Finally, the monitor loop is designed to make the entire program run with as much automation and safety as possible. It is always on and monitors every possible ACU reading from the limit switches to the load cells. It tracks for errors and makes sure that all readings are consistent with each other at any time (e.g. making sure that the turntable doesn't move while a source is deployed). If anything is wrong, it preemptively stops all motors, commands, and scripts, while looking into the problem in more details. It then proposes a solution before returning to its default state. In case of grave problems, it requests for an expert to be called. While a software can only be as perfect as the people writing it, this makes it in practice impossible for human error to cause damage to the capsules, cables, or other ACU equipment, during operations.

5.9 Conclusion and outlook

This chapter has presented the goals of OSIRIS' calibration and the work I have done to ensure its success. The ACU was presented, which is at the time of writing already at the JUNO surface sitting right next to the clean room in which it will be prepared soon before being transported and mounted on OSIRIS. The laser system, managed entirely by Tobias Sterr[24], was also presented. It will be able to calibrate the time offsets between PMTs to an accuracy of 25 ps and perform a solid charge calibration. The LED system will be able to cross-check these results with an accuracy of about 0.5 ns - but its greatest significance is that it can be used as soon as the PMTs are online. Indeed, since the LED capsule can be deployed inside of the acrylic vessel, it can be used even before the liquid scintillator or even the preceding ultra-pure water fills the vessel. However, as can be seen in section 5.4, there are difficulties in the time calibration related to reflections on the inner walls of the acrylic vessel. By using different measurements and implementing a new mode in the simulation software, I could identify LED-capsule positions for each PMT at which the levels of reflection-induced biases were similar. By combining data from these positions, one can therefore minimise the "damage" caused by these problems so that we can be confident in our results. The energy calibration, starting as soon as the detector is filled with liquid scintillator, was also discussed. I optimised the choice of isotopes and activities, taking into account measurement times, pile-up, and reconstruction capabilities, according to a realistic calibration plan. This

[24] Eberhard Karls Universität Tübingen

will allow us to measure the energy scale of the detector very accurately. To reduce the distortions in the energy reconstruction due to various effects that cause low-energy tails in the calibration peaks, I developed a method tailored to our isotopes to extract the full-absorption peaks with a precision to below one percent despite these distortions. Finally, I investigated the sensitivity of a capsule filled with natural potassium (i.e. a source with low ^{40}K activity) to quickly detect possible changes over time in the properties of the liquid scintillator. An investigation into the χ^2-based method I developed here showed that even a one percent change in the light yield could be detected within a few hours with a reasonable rate of type I and type II errors. This is especially important in OSIRIS through which new liquid scintillator will continuously be flowing in during normal operations, and will be of great help to the people working on the purification of the scintillator.

6 OSIRIS sensitivity to backgrounds from JUNO's LS

The main goal of the OSIRIS experiment is to verify the radiopurity of JUNO's liquid scintillator before it enters JUNO's central detector [207]. This is mainly achievable by measuring the contamination of ^{238}U and ^{232}Th in the LS by measuring the rates of fast beta-alpha coincidences of ^{212}Bi-Po and ^{214}Bi-Po decays, respectively [208]. Bi-Po coincidences occur in both ^{238}U and ^{232}Th chains where a bismuth beta-decay is immediately followed by a polonium alpha-decay, thereby leaving a very visible signal in the detector. Specifically, OSIRIS was designed to be sensitive to at least JUNO's *IBD* threshold of $1 \times 10^{-15}\,\mathrm{g\,g^{-1}}$. This was defined by the JUNO collaboration as the maximum ^{238}U and ^{232}Th contamination that would still allow JUNO's MH measurement to be possible [92]. Other requirements were set for the *solar* analysis (of ^{8}B neutrinos) at $1 \times 10^{-17}\,\mathrm{g\,g^{-1}}$ to $1 \times 10^{-16}\,\mathrm{g\,g^{-1}}$ - though as will be seen below these are unlikely to be reached by OSIRIS. Section 6.1 enumerates the background contributions to the analysis, section 6.2 outlines how the data is selected. Then section 6.3 explains how OSIRIS will perform these measurements and what sensitivity it can be expected to reach, while outlining my role in the calculations. Finally, section 6.5 explains that additional sensitivity is also reachable for ^{14}C, ^{210}Po, ^{85}Kr, and indirectly for ^{39}Ar as well. Determining rates or upper limits on all of these values will be helpful to future JUNO analyses as well, especially in the solar neutrino sector. The analyses performed in this section are based on Ref. [208], with small corrections but also key differences when it comes to the calculation of the sensitivity. The results were also published in the European Physical Journal C [207].

6.1 Expected backgrounds

6.1.1 Internal radioactivity

The first obvious source of radioactivity in OSIRIS are so-called "internals". These background events stem from radioactive isotopes that are still present in the liquid scintillator after the purification process, and whose decay energy is within OSIRIS' main energy range, *i.e.* roughly from 0 MeV to 3.5 MeV. All isotopes and their contamination values used in this thesis are based on Ref. [208]. The sources of internal background with the highest rates in OSIRIS are ^{14}C, ^{210}Po, and ^{210}Bi. However, when including the entire decay chains of ^{238}U and ^{232}Th, there is a total of 31 contributing isotopes. All of these isotopes are considered in the following analysis with expected rates based on the *JUNO solar* requirement, as defined in JUNO's yellow book [92]. Additionally, since Borexino experienced that not all measured isotopes are in secular equilibrium [181], a

factor of 10 was applied to artificially increase the expected rates of ^{210}Bi, ^{210}Po, and ^{85}Kr.

6.1.2 External backgrounds

The term *external backgrounds* refers to all of the background events stemming from radioactive decays originating outside of the liquid scintillator whose decay products still manage to reach the LS volume. Since α and β decays can in general not reach the LS, *externals* mostly only consist of gammas, except for those isotopes from the AV itself, where of course all decay products have to be taken into account. In general, the isotopes whose decay products are expected to reach the LS are ^{40}K, ^{214}Bi from the ^{238}U-chain, and ^{208}Tl from the ^{232}Th-chain. These can all stem from the AV, the water shield, the PMT glass, the PMT frame, the steel tank, and finally the rock walls of the experimental hall. The contamination of the different volumes were taken from a JUNO-internal database which contains radioactivity measurements of similar materials. For the water shield, measurements from Borexino's CTF [257] were used instead. Using those numbers (see Table 7.1 in [208] for more details), the main sources of externals are gammas from the rock surrounding OSIRIS and gammas from the PMT glass. A method called *geometrical biasing* was implemented in the simulation - by Christoph Genster[25] - which allows to precisely simulate external background in OSIRIS, overcoming the $\mathcal{O}(10^{10})$ attenuation of externals without having to generate untenable amounts of events [208].

6.1.3 Cosmogenic background

The last category of backgrounds in OSIRIS are so-called cosmogenic isotopes. These isotopes originate in spallation processes of cosmic muons on carbon (usually ^{12}C) in the scintillator. This produces many isotopes that can be easily rejected using veto cuts around muon events, but most notably it produces ^{10}C and ^{11}C whose mean lifetimes, about 20 s and 20 min, respectively, vastly exceed the time in which parts of the detector can be vetoed. They therefore constitute a possibly large source of background in every liquid scintillator detector, though in OSIRIS the expected rates of cosmogenic background are fairly insignificant when compared to those of externals.

[25] IKP-2, Forschungszentrum Jülich

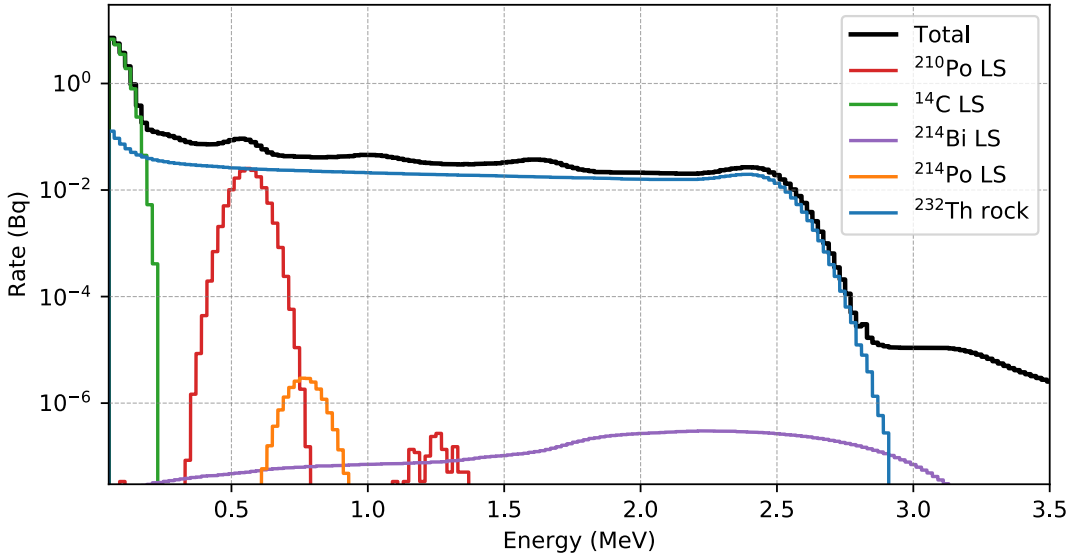


FIGURE 55: Energy spectra of the main background contributions in OSIRIS, as well as an example ^{214}Bi-Po signal assuming a ^{238}U and ^{232}Th contamination of 10^{-15}, in a 150 cm FV. As one can see, the main background contribution is from ^{232}Th from the rock surrounding OSIRIS, and from ^{14}C and ^{210}Po in the LS. The extra bump from ^{210}Po events near 1.3 MeV is from a rare decay mode.

6.2 Bi-Po selection

A series of cuts were performed in order to identify as many Bi-Po coincidences as possible while of course rejecting as much background as possible. In order to study OSIRIS's sensitivity to ^{238}U and ^{232}Th contamination, these cuts were tuned on simulated background events. Because at this point in time many of the reconstruction algorithms for OSIRIS don't exist, this analysis does not go over the data event-by-event, focusing instead on a purely statistical calculation of the sensitivity. This is done by using estimated reconstruction performances and energy and position spectra instead, used to evaluate the performances of the different cuts. In order to create the spectra used in the evaluation of the cuts, reconstruction performances of 6%@1 MeV on the energy and of 14 cm@1 MeV were assumed.

The first cut is a fiducial volume cut with, for simplicity, a single value which is then applied to both the z and ρ coordinates in a cylindrical coordinate system. This makes sense because we expect most background events to occur near the outer edges of the LS volume. Energy spectra created assuming a FV of 150 cm can be seen in fig. 55. The second cut is an energy cut which is applied separately for the **prompt window**, meant to select $^{212-214}$Bi events, and for the **delayed window**, meant to select $^{212-214}$Po events. As such, in the following, all events with an energy falling in the prompt window

are referred to as a bismuth candidates, and all events with an energy falling in the delayed window as polonium candidates.

The next cut is on the time delay between the prompt and the delayed candidate. Thanks to the very low mean lifetime of $^{212(214)}$Po, of 431 ns(237 µs), this is the most powerful cut. The final cut is on the reconstructed distance between the prompt and the delayed events. The reasoning is that since the polonium isotope decays in much less time than is required to travel any significant distance in the scintillator for the atoms, one can assume that they occur at the same position. The distance cut is also a very powerful cut.

The signal and background efficiencies of all cuts (defined as *signal efficiency*: how much signal is kept, *background rejection efficiency*: how much background is rejected) were then calculated, as previously mentioned, on a statistical basis using scaled expected spectra instead of an event-by-event basis. The efficiencies of the fiducial volume cut are not clearly defined because of their correlation with other cuts. Energy and position spectra were created instead for a given fiducial cut value and the subsequent efficiencies calculated with respect to these spectra. The efficiencies of the energy cut were simply determined by integrating the spectra in the defined energy windows. Afterwards, the efficiencies of the time cut were calculated using the following equations:

$$\begin{aligned} \epsilon_t &= e^{-t_0/\tau_{Po}} - e^{-t_1/\tau_{Po}}, \\ \epsilon_t' &= 1 - e^{-R_{Po}\cdot(t_1-t_0)}, \end{aligned} \tag{55}$$

where ϵ_t and ϵ_t' are the signal and background rejection efficiencies of the time cut, respectively, t_0 and t_1 are the starting and end times of the delayed time window in which events are accepted, τ_{Po} is the mean lifetime of the polonium isotope, and R_{Po} is the rate of the polonium isotopes after FV and energy cuts. For the determination of the efficiencies on the distance cuts, a more complicated algorithm was developed which is entirely based on Ref. [208]. It unfortunately had to be recreated from scratch because of hardware issues, but the results agree well with Christoph Genster's. As such, the used cut values, which were optimised for my sensitivity calculation described in section 6.3, are mostly identical to Ref. [208] with only small changes. They are shown in Table 18 for completeness.

6.3 Expected sensitivity to ^{238}U and ^{232}Th

Once the cuts described in the previous sections were applied, I calculated the number of events expected from background only. Using this number, it is possible to calculate the sensitivity using the *counting experiment* strategy described in [227]. The sensitivity is

Cut		^{212}Bi-Po Values	^{212}Bi-Po Efficiencies	^{214}Bi-Po Values	^{214}Bi-Po Efficiencies
FV	ρ	150 cm	-	145 cm	-
	$\lvert z\rvert$	150 cm	-	145 cm	-
Energy	E_{Bi}	0.14 to 2.50 MeV	ϵ_{Bi}: 0.98	1.4 to 3.5 MeV	ϵ_{Bi}: 0.82
	E_{Po}	0.86 to 1.5 MeV	ϵ_{Po}: 1.00	0.65 to 1.1 MeV	ϵ_{Po}: 0.99
Time	t	0.2 to 2 µs	ϵ_t: 0.62 $\bar{\epsilon}_t : 2.1 \cdot 10^{-6}$	0.2 to 711 µs	ϵ_t: 0.95 $\bar{\epsilon}_t : 6.8 \cdot 10^{-4}$
Distance	r	72.5 cm	ϵ_r: 0.98 $\bar{\epsilon}_r$: 0.05	45 cm	ϵ_t: 0.88 $\bar{\epsilon}_t$: 0.02
Total selection efficiency			0.59		0.67

TABLE 18: Cut values and their different resulting uncertainties. ϵ_X is the signal efficiency of cut X, and $\bar{\epsilon}_X$ represents how much of the background is left after a cut. It is equal (see text) to $1 - \epsilon'_X$.

defined as a number N_S: the median signal strength needed in order to be able to reject the background-only hypothesis with a p-value of $1 - \alpha$. An easier calculation to follow, which outputs almost the same values as the one used, is a Poisson based sensitivity: for a given expected number of background events N_b, one calculates the median of $P_{N_b}(x)$ (where $P_\lambda(x)$ is a Poisson distribution with an expected value of λ), and then finds an N_S such that the integral of $P_{N_S+N_b}(x)$ from said median to infinity is equal to α. The actual calculation used considers additional effects such as possible under-fluctuations and possible control measurements. The result of those calculations (using $\alpha = 0.9$) can be seen in fig. 56, which was made by myself and published in [207]. It is important to note that this plot is only valid for a continuous measurement of a single LS sample over time. As one can see from the figure, the JUNO IBD threshold can be reached in a few hours, however several days are needed to reach the JUNO solar level. For low times, both curves have the same slope - because they are both in a *no-background* regime where the sensitivity only depends on measurement time. The ^{212}Bi-Po curve is slightly higher because the signal efficiency for ^{212}Bi-Po is smaller than for ^{214}Bi-Po. However, differences in the background rates make the ^{214}Bi-Po exit the *no-background* regime sooner which leads to better ^{212}Bi-Po sensitivity on timescales of more than one day.

6.4 Batch- vs Continuous-mode operation

OSIRIS will be operated in two different modes. In the *batch* mode, LS will reside in the AV for several days or weeks. This mode will be adopted during the commissioning phase of JUNO's purification plants. Then, during the actual filling of the JUNO central detector, the *continuous* mode will be adopted where LS will continuously flow through OSIRIS at a rate of about one cubic meter per hour, which corresponds to the entire 18-ton LS volume of OSIRIS being exchanged in one day. This represents about 18% of

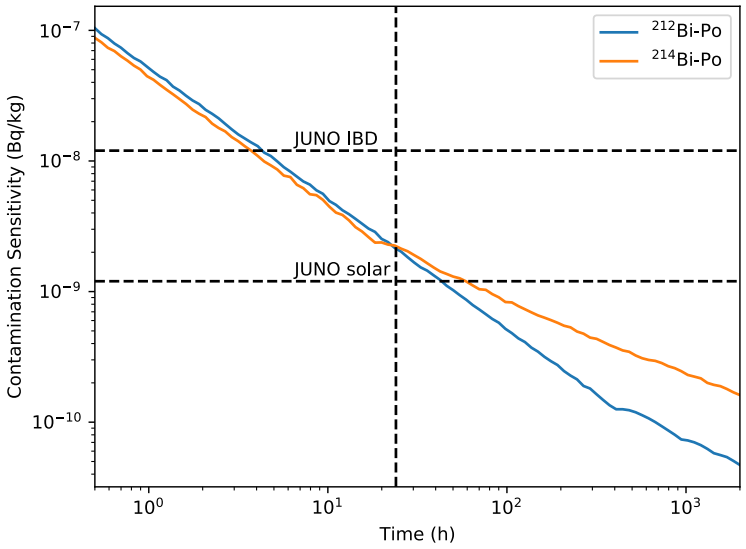


FIGURE 56: Sensitivity to $^{212(214)}$Bi-Po contamination in OSIRIS as a function of measurement time. $^{212(214)}$Bi-Po contamination can be translated directly to a contamination on ^{232}Th(^{238}U). As one can see the JUNO IBD threshold can be reached in a few hours, however several days are needed to reach the JUNO solar level. The fluctuations in the curves reflect the uncertainty from the Toy-MC method used in the calculation (see text). The vertical dashed line corresponds to a measurement time of one day. This figure was published in [207].

JUNO's filling capacity. While the *continuous* mode has of course less sensitivity than the *batch* one, it can still be used to provide fast alarms and even vetoes if need be. More details can be found in Ref. [207].

At this point it is important to mention that an initial radon contamination is to be expected from the pipe system. This consists of ^{220}Rn and ^{222}Rn, whose decay products contain ^{212}Bi-Po and ^{214}Bi-Po, respectively. This initial contamination is expected to be, from internal JUNO estimations, between $130\,\mathrm{d}^{-1}$ to $830\,\mathrm{d}^{-1}$. Although this of course always worsens the aforementioned sensitivity, this is not a huge problem in the *batch* mode as it is possible to fit the decaying component. In practice, this affects the ^{238}U sensitivity much more than the ^{232}Th one because the half-life of ^{220}Rn is only about one minute (while the half-life of ^{222}Rn is about 3.8 d).

The most important case is however that of the *continuous* mode since this is the one that will have to be used during the actual filling of JUNO. One cannot reliably perform a time-fit here because the entire LS volume is exchanged every day. It is however possible to measure the event rate as a function of the z-coordinate, because the scintillator flows strictly from the top to the bottom of the AV. After one day of measurement, this, however, still significantly worsens the sensitivity (compared to the no-radon case) to between $6\times10^{-15}\,\mathrm{g\,g^{-1}}$ to $1.4\times10^{-14}\,\mathrm{g\,g^{-1}}$ [207] (reminder: the JUNO

IBD requirements are $1 \times 10^{-15}\,\mathrm{g\,g^{-1}}$), depending on the initial radon contamination. One can however influence the results of the fit by adding a pull-term on the initial radon rate. As such, if the initial radon contamination is known with a precision of 10%, one reaches the JUNO IBD requirement within one day. While it remains to be seen how precisely this will be measured in OSIRIS, this should in principle be achievable through dedicated long-term *batch* measurements, assuming that the rate of in-flowing radon stays stable.

6.5 Sensitivity to other LS contaminants

^{85}Kr and ^{39}Ar contamination ^{85}Kr is a beta emitter that is a particularly nefarious background for JUNO's ^{7}Be neutrino analysis, see section 4.3. The isotope is present in high quantities in the atmosphere, globally around 5000 PBq, originating mostly from nuclear fuel reprocessing [241]. Consequently, even small exposures to air can introduce relatively large contamination in the liquid scintillator. Fortunately, while ^{85}Kr usually decays directly to stable ^{85}Rb, with a branching ratio of 99.57%, it in rare cases (about 0.43%) also undergoes a beta decay to a meta-stable ^{85m}Rb isotope, with an endpoint energy of 173 keV. The latter is immediately followed by a de-excitation with a gamma emission of 514 keV[26].

This fast coincidence is exploitable in the same way as the Bi-Po coincidences. Using nearly the same methodology as the one described above, Philipp Kampmann[27] and I investigated OSIRIS' sensitivity to ^{85}Kr. One key difference is that the energy spectrum of the prompt event goes all the way down to zero. This reaches a range which overlaps with ^{14}C events and where the energy dependence on the trigger efficiency of OSIRIS must be taken into account. This efficiency was calculated to be 69.6%, by Philipp Kampmann using a full simulation chain including the expected performance of the PMT electronics in OSIRIS, over the entire prompt spectrum (which is inside of the cut boundaries in this analysis). The total efficiency of the delayed event is about 96% for its total spectrum, but since in this case only events above 300 keV are selected its trigger efficiency in this analysis is 100%. The values of the different cuts (see section 6.2) were optimised using a grid search and can be found in Table 19. It is worth noting that this efficiency can be increased if a pre-trigger DAQ-window of 5 µs is introduced. In this case, a trigger will in all cases be issued by the delayed gamma event, permitting to find a sub-threshold prompt signal by a backward-search in the pre-trigger window [207].

Based on these cuts and the same counting-experiment statistical treatment used for the Bi-Po analysis (see section 6.3), one obtains a sensitivity of about $1 \times 10^5\,\mathrm{d^{-1}\,kt^{-1}}$ at 90%

[26] https://www.cdc.gov/niosh/ocas/pdfs/tbd/pine4-r1.pdf
[27] IKP-2, Forschungszentrum Jülich

C.L. within a day of measurement. This sensitivity is improved to $6 \times 10^3\,\mathrm{d}^{-1}\,\mathrm{kt}^{-1}$ after six months of measurement, which is still one order of magnitude above the $500\,\mathrm{d}^{-1}\,\mathrm{kt}^{-1}$ purity goal set by the JUNO collaboration [92]. While OSIRIS will thus not be able to meaningfully measure rates of ^{85}Kr in the amount that they are expected in JUNO, the obtained sensitivity is still extremely useful in the case of air-leaks. Since ^{85}Kr is so characteristic of an air-leak and because of the high expected rates in that case, it is clear from this study that we will be able to very quickly identify that an air leak has occurred. Because different sections of the pipe system surrounding OSIRIS can easily be cut-off and thus isolated from one another, ^{85}Kr identification is thus expected to help in not-only identifying that an air-leak has occurred, but also to pinpoint its location in the pipe system. Additionally, assuming that the ratio between ^{39}Ar and ^{85}Kr in OSIRIS is the same as in the atmosphere (about a factor 1:100), any measurement or upper limit on ^{85}Kr can in principle also be used as a measurement or limit on ^{39}Ar. After a measurement time of six months, this is enough to constrain the rate of ^{39}Ar down to a level that would be insignificant for JUNO's solar analysis [207].

Detecting ^{14}C and ^{210}Po While all of the signals until now are so low that sensitivity calculations needed to be performed to see if they can even be measured, there are two isotopes whose decays are expected to contribute in large amounts to the energy spectrum of OSIRIS and with very sharp spectral shapes: ^{14}C and ^{210}Po. These components can be measured through a spectral fit of the singles spectrum, see fig. 55. The ^{210}Po peak is well recognisable and OSIRIS's sensitivity to it is $1 \times 10^{-24}\,\mathrm{g\,g}^{-1}$ (95% C.L.) in a day [207]. While a pulse shape discrimination could improve this result, it

TABLE 19: Efficiencies of the different data selection cuts for the ^{85}Kr measurement. As in the Bi-Po selection, Δt and Δr represent time and distance cuts. For the latter, the numbers in brackets represent the rejection efficiency for accidental coincidences. In addition to the efficiencies of the selection cuts, the trigger efficiency is also considered in the total detection efficiency. The efficiency on the energy cut of the prompt event is 100% because it includes the entire prompt spectrum, however a trigger efficiency of 70%, shown as "Event trigger" must also be taken into account for it.

	$^{85}Kr - Rb$	
Cut	Values	Efficiency (%)
FV	<150 cm	-
E_{prompt}	<0.2 MeV	100
E_{delayed}	0.36 MeV to 0.50 MeV	83
Δt	0.35 µs to 4.76 µs	74
Δr	<150 cm	84 (63)
Event trigger	5 hits in 70 ns	70
Total		36

already corresponds to rates expected in JUNO [92]. In *batch* mode, OSIRIS is expected to measure ^{210}Po with a precision of 5%. For ^{14}C the precision reaches less than 1%, although systematic uncertainties, which are not taken into account here, will probably dominate the final uncertainty.

6.6 Conclusion and outlook

During both the commissioning of the purification system and during JUNO's filling phase, OSIRIS will be able to give live feedback to said system by measuring fast Bi-Po coincidences. OSIRIS will also be able to quickly detect air-leaks by exploiting a rare ^{85}Kr decay chain, and measure additional $^{220-222}Rn$ coming from the emanation and possible air contact during purification. It is worth mentioning that the sensitivity achieved by OSIRIS strongly depends on the levels and stability of said background $^{220-222}$Rn contamination coming from the LS as the resulting Bi-Po events would be indistinguishable from ^{232}Th or ^{238}U signals. This is especially relevant in the *continuous* mode as the relatively short decay times of radon (5.5 days for ^{222}Rn) let most of it decay when using the *batch* mode. Under reasonable assumptions for the initial $^{220-222}$Rn rates in OSIRIS, the measurements will be able to guarantee JUNO's IBD-requirement of a $1 \times 10^{-15}\,\mathrm{g\,g^{-1}}$ contamination of ^{232}Th and ^{238}U in the liquid scintillator in a *continuous* measurement of flowing scintillator, while JUNO's solar requirements $1 \times 10^{-17}\,\mathrm{g\,g^{-1}}$ to $1 \times 10^{-16}\,\mathrm{g\,g^{-1}}$ will be able to be probed in some *batch* mode measurements.

The OSIRIS setup will continue to be useful after the filling of JUNO. In particular, Serrapis (SEarch for RAre PP-neutrinos In Scintillator) is currently under investigation [258]. Serappis is an upgrade of OSIRIS in which the PMTs are assembled spherically around the acrylic vessel and are equipped with Winston cones - thereby reaching 1500 p.e./MeV. With additional shielding against external backgrounds (additional $\approx$ 1.2 m.w.e. of e.g. concrete blocks) and by placing the PMTs at a greater distance from the scintillator, one could achieve external background levels of 2.8 mBq in a fiducial volume of 1.2 m. With extremely pure scintillator (^{14}C contamination of less than $1 \times 10^{-20}\,\mathrm{g\,g^{-1}}$), one could then reach a precision of 3.4% on pp-neutrinos (compared to roughly 10% from Borexino [46]) within roughly four years of data-taking. Obtaining such a low-^{14}C LS sample is however expected to be very difficult. These estimations were aided by the work performed in this thesis, particularly in the estimation of external backgrounds for OSIRIS.

Conclusion & Outlook

CNO neutrinos are neutrinos from a rare fusion sub-process in the Sun called the CNO cycle. Section 3 presented a modern Bayesian analysis framework which I designed to use spline functions to approximate the shape and size of a region inside of the Borexino detector in which polonium decays are rarest. Polonium in Phase-III of Borexino's data-taking comes from two sources: a ^{210}Pb-induced *scintillator* component that is in equilibrium with ^{210}Bi decays, and *vessel* polonium that detaches from the inner nylon vessel before entering the fiducial volume through convection. Finding an upper limit on the total Polonium rate can thus deliver an upper limit on the bismuth that is in equilibrium with polonium (assuming that the rate of bismuth decays in the fiducial volume is otherwise constant, which was found to be true). Before 2016, polonium in the detector could travel large distances through thermally induced convective currents. Since then a thermal insulation of the detector has all but stopped these movements. Hence, a modelling of the spatial distribution of polonium decays (which can be efficiently marked through their alpha decays) could provide an upper limit on the rate of *scintillator* polonium - which can be turned into an upper limit on ^{210}Bi decays. This constraint, which can easily be determined using my model - implemented by Sindhujha Kumaran and myself - was able to deliver a 5σ evidence for the existence of CNO neutrinos, whose interaction rate in Borexino was found to be $R(\mathrm{CNO}) = 7.2^{+3.0}_{-1.7}$cpd/100t. Additionally, the framework, which incidentally calculates the so-called Bayesian evidence, can be used to test different models against the data.

In the context of sensitivity studies for the measurement of solar neutrinos with JUNO, I have designed and developed a statistical software tool, together with Anita Meraviglia and Luca Pelicci, which is still being used today to efficiently create large amounts of on-the-fly Poisson-fluctuated toy-datasets on which spectral fits are performed (see section 4). The software is completely modular and configurable with external files, so that any modification is quite straightforward and such that any modification to the tested scenario is easy to implement. It was also built with the possibility of being used in other experiments as a long-term goal, and though minor modifications were needed, it is now also being used in the context of JUNO antineutrino analyses. Using the software, it was shown that JUNO will achieve unprecedented measurements of ^{7}Be ($\leq 1\%$ precision in all studied scenarios), but it was also shown that a large ^{85}Kr contamination, which could rise to very high levels following a simple air-leak or emanation from the acrylic vessel, would destroy JUNO's capabilities. For *pep* neutrinos, the situation is more difficult, though the results also depend a lot on the tagging power of the Three-Fold-Coincidence (TFC) algorithm that will be used in JUNO (TFC algorithms veto detector regions containing ^{11}C events after coincidences between the parent muon and a neutron

capture). Under reasonable expectations for the radiopurity of the scintillator however, it was shown that JUNO will exceed Borexino's precision after about one year of data-taking, except in the so-called *IBD* scenario. This is a worst-case radiopurity scenario in which JUNO will only barely reach the requirements set by JUNO's IBD analysis - and in which unprecedentedly precise *pep* measurements with JUNO are not guaranteed. Finally, it was shown that JUNO will be able to reliably measure CNO neutrinos only if strict radiopurity scenarios are achieved. Should they be, then JUNO is expected to surpass Borexino's precision in six years, even without a constraint on ^{210}Bi. When using an external constraint on *pep* neutrinos, then JUNO is expected to surpass Borexino's precision in only four years in all but the aforementioned worst-case scenario - in which case surpassing Borexino's results will not be possible. In its planned ten years of data-taking, JUNO will so be able to reach about a 10% precision on CNO. Preliminary results also indicate that JUNO will be able to separate the ^{13}N and ^{15}O contributions in the CNO spectrum with 10% to 30% uncertainty. All-in-all, while large uncertainties exist on the expected results because of the unknowns in JUNO's achieved radiopurity, JUNO's immense statistics were shown to have the potential for unprecedented precision in solar neutrino measurements. However, the analysis presented in this thesis does not consider systematic uncertainties. In the future, a more detailed treatment of low-energy trigger simulations could even lead to a measurement of *pp*-neutrinos. Since the JUNO collaboration also showed that the measurement of ^{8}B neutrino will be possible, this would make JUNO the second experiment to ever measure all neutrinos from the pp-chain.

A large part of the work in this thesis went into the development of the source calibration program of OSIRIS, as shown in section 5. This program encompasses all aspects of calibration by using the Automated Calibration Unit: a machine capable of delivering three different capsules directly inside of the liquid scintillator volume. These three capsules contain a three-isotope gamma source (^{137}Cs, ^{65}Zn, and ^{60}Co), a low-activity ^{40}K source, and a fast pulsed LED, respectively. It was shown that the former could, when using specially constructed physically-motivated fit functions, be used to reconstruct the energy scale of the detector with a high fidelity. The second was shown to be able to provide fast warnings in case of even percent-level changes in the properties of the liquid scintillator (or of other elements of data-taking for that matter) - and a method was developed with which one can tune the performances of said change detection while minimising the time needed until the alarm is triggered. The last capsule was shown to enable a calibration of the time-offsets between the PMTs with a precision of 0.1 ns and a bias of less than ± 0.2 ns, which is well below the innate TTS expected from the PMTs (around 1.2 ns). This involved several series of measurements both in Tübingen and in Aachen. In the context of the calibration studies, a LabView control software,

based on old Daya Bay software, was built nearly from scratch to ensure the safety of the liquid scintillator and of the ACU at all times - and to make sure that the ACU doesn't need to be opened any more than is strictly necessary. The installation of the ACU on OSIRIS was also prepared in many details, which due to the pandemic will have to be at least partly remote. In the near future, the LabView software will be updated by Cornelius Vollbrecht to include communication with the run-control of OSIRIS, as well as to provide a more user-friendly interface. Around August-September 2022 the installation of the ACU on OSIRIS will begin, as will the first measurements taken with the LED capsule. The radioactive sources, which are in the process of being ordered, are expected to be inserted around October-November. After an lengthy initial calibration to make sure that everything is understood, plans for weekly calibrations were made in the context of this thesis. Because one should always expect the unexpected once the experiment actually starts, programs written during this thesis were already shared and modified to allow for easy recalculations if needed.

In order to adapt OSIRIS's main analysis of detecting fast Bi-Po coincidences to detector updates, I developed a framework parallel to that developed by Christoph Genster (see section 6). By doing so, the analysis was also refined to a more robust definition of sensitivity. Those results were subsequently optimised, resulting in slightly more optimistic values as well as a more realistic behaviour of sensitivity with measurement times which take OSIRIS's exit of the "no-background mode" into account - with respect to existing calculations. As such it was confirmed that in a naive analysis OSIRIS will be sensitive to ^{238}U and ^{232}Th to the level of $1 \times 10^{-15}\,\mathrm{g\,g^{-1}}$ within a few hours, and to levels of $1 \times 10^{-16}\,\mathrm{g\,g^{-1}}$ within a few days. Given a more realistic scenario featuring initial radon contaminations, the situation worsens and OSIRIS appears not to be able to be sensitive to $1 \times 10^{-15}\,\mathrm{g\,g^{-1}}$ within a day - which is the time it takes for liquid scintillator to traverse the detector. However, due to $^{220-222}$Rn's low half-lives, it was shown by Michael Wurm that the z-dependence of Radon signals could be used to extract the underlying ^{238}U and ^{232}Th contributions with sufficient precision within one day - should the rate of initial radon be constrained to better than 10%. This is a realistic requirement to extract from initial batch measurements, at least if the rate of the radon contamination does not significantly change over time. My calculations for Bi-Po sensitivity were then also applied to the search of a rare ^{85}Kr decay resulting in a similar fast coincidence. This was used to show that while OSIRIS is not sensitive to any expected levels of ^{85}Kr, it will be able to use it to identify possible air-leaks. In the context of solar neutrino measurements with JUNO - OSIRIS will thus only be sensitive to the IBD scenario for ^{238}U and ^{232}Th. For ^{85}Kr however, OSIRIS will be sensitive enough to limit the uncertainty on ^{7}Be neutrinos to less than about 2% in one year (from the effects of ^{85}Kr only) - though only for ^{85}Kr that was already present in the scintillator at the OSIRIS level,

this doesn't help in case of further air-leaks or vessel emanation. In the future, once OSIRIS is equipped with a robust vertex and waveform reconstruction, these studies can be repeated without the approximations that I used to get less conservative results.

Appendices

Appendices

A Borexino and the ^{14}C energy region

Previous comprehensive analyses by Borexino performed the multivariate fit in a region starting at about 190 keV [46]. Since the electron recoil spectrum of CNO neutrinos extends down to very low energies, it is natural to try and include as much of the energy range as possible. As for every liquid scintillator experiment, the analyses meet a hard limit at the energy region of ^{14}C, whose decay rate is comparatively extremely high, even in highly radio-pure environments like Borexino's core. The reason for the difficulty is not only the high rate itself, but also its effects on the measurement procedures in the form of pile-up.

In particle physics, "*pile-up*" refers to recorded events caused by multiple physical processes happening so close in time that it becomes very difficult to (at least correctly) distinguish between them. In cases where event rates are high and measurement time windows are small, this effect can have a large influence on results. In liquid scintillator experiments, the energy of an event is oftentimes estimated by counting *e.g.* the number of PMTs which are hit in a given time window after a trigger. In this case, the most common kind of *pile-up* occurs when the light of two different events reach the PMTs nearly simultaneously. Instead of correctly measuring two separate events, one will instead measure a single event with an erroneously large energy. The event will also typically be reconstructed at a very incorrect location.

In the very-low-energy region of the solar neutrino analysis of the Borexino experiment, pile-up is one of the most critical background components [183]. Using $r_p = r_{^{14}C}^2 \cdot \Delta T$ with $\Delta T = 230$ ns and a FV of 280 t, and the rate $r_{^{14}C}$ of ^{14}C decays in Borexino being about 40 Bq/100t, one can quickly approximate the rate of ^{14}C-^{14}C pile-up to be of about 90 cpd/100t. While other effects will influence this rate in a more exact calculation, this order of magnitude is still helpful to know. Because of similarities in rates and spectral shapes between pile-up and *e.g.* the *pp*-neutrino component (around 130 cpd/100t), it is crucial to precisely know not only said spectral shape of pile-up events, but also its occurrence rate.

Following naturally from the two analysis methods used in the spectral fits of Borexino's solar analyses (analytical fit and MC fit), two main methods were used to estimate pile-up characteristics in the experiment. The methods derived their estimations from data and from MC simulations, respectively. Both of these analysis methods were used in parallel, and their results compared. It was shown on phase-II datasets that both methods gave compatible results when used over the whole energy range. Appendices

A.1, A.2, A.3 describe both methods surrounding pile-up as part of the efforts made to extend the energy region of the CNO analysis. This is done mainly for academic reasons since, as Appendix A.4 explains, the low-energy regions ended up having to be discarded in the final analysis.

A.1 Data driven pile-up

While the CNO analysis was performed using only the Monte-Carlo fit, this section describes how pile-up was treated in the analytical fit used in previous analyses. This nevertheless allows for a good cross-check of the MC methods, including of the simulation itself, and provides relevant insights as explained below.

In the analytical fit of the Borexino solar neutrino analyses, one can use the "*npmts_dt1*" variable to construct a pile-up spectrum from data. The variable is defined as the number of hit PMTs in a fixed-length 230 ns time window. To do so, a 230 ns time window is chosen randomly inside of the last 6 µs of a (16 µs) data-acquisition gate, and overlapped with the original event time window (also of 6 µs). The resulting event is then reconstructed in the same way as any other data event and subsequently compared to the original, unperturbed, event. A schematic representation of this procedure can be seen in Figure 57.

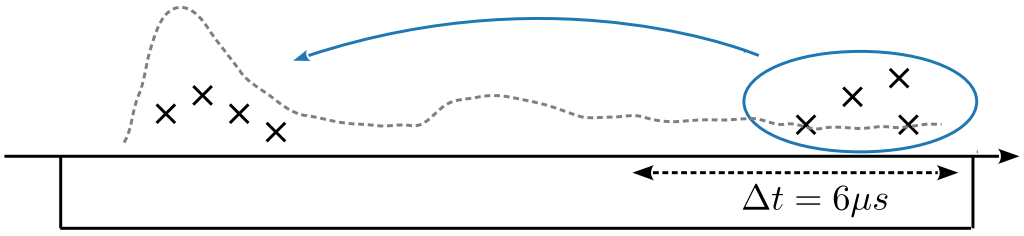


FIGURE 57: A schematic representation of the synthetic pile-up procedure. The dashed line shows the probability of an event in the DAQ gate. Credit: *Mariia Redchuk [189]*

This so-called "synthetic pile-up" is a relatively simple way to estimate the rate and spectral shape of pile-up events as it only relies on two assumptions: that the tails of the events don't significantly affect the last 6 µs, and that the probability of an event occurring in the second time window is identical to that in the first. The former was verified to be true and the latter is expected from radioactive decays. However, there are a few systematic effects which limit its power. Firstly, by definition, dark noise from the PMTs is accounted for twice with the overlap of the time windows. Secondly, this can only account for events in which the first component has already triggered on its own, while in general it is also possible for two events which would not have triggered by themselves to form one pile-up event. To avoid labelling double-counted dark noise

as pile-up, a variable called E_{min} was introduced. Overlapped events were only selected to be "synthetic pile-up" if the reconstructed energy of the perturbed event differed by at least E_{min} to that of the original event. This value must unfortunately be chosen somewhat arbitrarily, which is a source of systematic uncertainty in the fit.

In order to account for the contribution of dark-noise in the analytical fit, a convolution method [183] was instead used in the energy range below E_{min}. This method consists of first acquiring random events (with no trigger) that consist of mostly dark noise and ^{14}C events. The spectra from different fit components are then convolved with the so-obtained spectrum, up to the value of E_{min}, after which the "synthetic pile-up" comes into effect. While it is important to understand the limitations of data-driven pile-up, one must also consider the great advantages it carries. For example, unlike with the convolution method where one doesn't have individual events, the effects from different analysis steps and selection cuts are automatically taken into account, as well as efficiencies from the different reconstruction algorithms. Furthermore, the absolute rate of these collected events is, by construction, known.

A.2 Monte-Carlo pile-up: Composition

In the MC simulation of the Borexino detector, it is possible to very precisely obtain pile-up spectra, meaning that once the principal components of the pile-up spectrum are known, then it is possible to simulate those events with no double contamination from *e.g.* dark noise or systematics from the choice of E_{min} (see above). Previous studies [259] on the subject have identified what those components are. The first component is the obvious ^{14}C-^{14}C pile-up, which is clear from simple order-of-magnitude estimations. Afterwards, more non-negligible components were revealed by studying second-cluster events. Second cluster events are detected clusters inside of a DAQ gate that was opened by another event, they are thus not affected by any trigger threshold. A lot of those events were shown to be affected by a fiducial volume cut, which suggested a strong contribution from externals to pile-up events. Indeed, extrapolating from radioactivity measurements inside of the FV [183] using curves of radial event distributions [259], and using measurements from different Borexino detector components [260], one can estimate the rate of external backgrounds (EB) in the entire SSS volume to be about 7 kBq. Most of those events are very small disturbances, with mostly only a few PMT hits per event, originating from γ's in the buffer with only partial energy deposition in the LS, where quenching also significantly reduces the measured energies. While those events would not trigger on their own, they can however overlap with a *e.g.* ^{14}C, altering its measured energy, thereby forming pile-up. The non-negligible external background components of pile-up were found to be ^{14}C-EB and ^{210}Po-EB.

A.3 Monte-Carlo pile-up: Rates

As explained in the previous section, to simulate the shape of the MC pile-up, three components must be simulated, followed by the determination of their relative rates. These rates are not a-priori known, not only because it is difficult to measure the rates of the individual particle species accurately, but also because of the various systematic effects and efficiencies involved in (pile-up) event reconstruction. Indeed, some scenarios are very different for pile-up when compared to simpler events. Two events which will constitute one pile-up event can in fact be several meters apart inside of the detector, even outside of the fiducial volume, and still be reconstructed as a valid event.

However, because ^{14}C is homogeneously distributed within the IV, and because its decay rate is precisely known from second cluster measurements (see Section 2.1), it is possible to precisely estimate the rate of ^{14}C-^{14}C pile-up. To perform this calculation, one must first simulate pairs of ^{14}C events that overlap in time. Since for random events, with a known mean occurrence rate, the probability distribution of the time δt between two such events can be precisely described with an exponential distribution, a Laplace distribution was used to sample the time interval between any two decays. While for the other pile-up simulations mentioned later in the thesis an asymmetric Laplace distribution was used, for the case of ^{14}C-^{14}C it was of course symmetric. The secondary events were placed within 2 µs of the respective primary ones (on either side).

The resulting events then underwent normal reconstruction steps, followed by a series of cuts to determine whether they actually constitute pile-up events. When simulating two consecutive events, ev_1 and ev_2, one obtains a ,,merged" ev_3. If a cluster was found in any event, a number of hits n_i was associated to it, where $i \subset [1, 2, 3]$. Using this nomenclature, the pile-up cuts were:

1. MC pile-up
   - If ev_1 has a cluster $\wedge$ $n_3 \neq n_1 \rightarrow$ pile-up
   - If ev_1 has no cluster $\wedge$ ev_2 has no cluster $\wedge$ ev_3 has a cluster $\rightarrow$ pile-up
   - If ev_1 has no cluster $\wedge$ ev_2 has a cluster $\wedge$ $n_3 \neq n_2 \rightarrow$ pile-up
2. Simulated synthetic pile-up
   - If ev_1 has a cluster $\wedge$ $n_3 - n_1 \geq E_{min} \rightarrow$ pile-up

The resulting events then underwent further normal solar analysis cuts, and the rate of ^{14}C-^{14}C pile-up was calculated using:

$$r_p = \frac{N_{ana}}{N_{sim}} \cdot \frac{\sum_i (M_{IV}^i)^2 \cdot L^i}{\sum_i L^i} \cdot r^2 \cdot \delta T \cdot \frac{864}{M_{FV}}, \tag{56}$$

where r_p is the rate of ^{14}C-^{14}C pile-up in cpd/100t, the subscript i denotes the id number of the simulated data run, N_{ana} is the number of events that survive all the cuts, N_{sim} is the number of simulated overlapped events, M_{IV} is the (time-dependent) mass of the scintillator in the IV, L^i is the live-time of the run (in seconds), r is the rate of ^{14}C in cpd/100t, δT is 230 ns, and M_{FV} is the mass of the fiducial volume. The masses are in units of 100 t.

Determining the relative rates of the pile-up components is a more complicated task. While alternative methods exist [259], their associated uncertainties are quite large. In this analysis, the relative rates were determined by fitting MC-generated pile-up spectra to the synthetic pile-up (see Appendix A.1). The synthetic pile-up spectrum already contains all possible pile-up components and includes the different cuts and efficiencies, but there are some systematic effects, especially at low energies. A straightforward way to avoid those effects is simply to fit only above an energy which is sufficiently above the value of E_{min} (see Appendix A.1). The results of a fit to the synthetic pile-up can be seen in Figure 58. The MC events used in this fit were generated from the same *raw* (pre-cut) sample of events as the events described above. They were selected using an alternative pile-up cut designed to mimic synthetic pile-up creation (see above). After the fit, one can relate the fitted rates to the MC pile-up rates by scaling to the number of events that survive the respective sets of cuts.

In principle, one can greatly improve the uncertainty on the relative rates by performing several fits with different values of E_{min}. However, as previously mentioned and for reasons outlined in the next section, the inclusion of the very-low-energy region in the fit had to be abandoned for the Phase III dataset, which is why the analysis stopped at the single fit stage.

A.4 Excluding pile-up from the CNO analysis energy range

The lifetime of any detector is limited by the instruments used in its design. In the case of Borexino, who already surpassed its original design goals, we observed a strong decrease in the number of active channels over time since the start of data-taking in 2007. As such, only about half of the $\approx$ 2000 Borexino PMTs were used in the measurements of the Phase-III data (spanning from July 2016 to January 2020). Consequently, the energy resolution, as well as the vertex reconstruction precision, got noticeably worse over time ($\sigma_E/E \approx 6\%$, $\sigma_{x,y,z} = 11$ cm for electron events at the detector centre at the time of writing [68]). This of course disproportionately affects the low energy regions.

It was found that applying analysis strategies from previous Borexino measurements did not produce reasonable results in the low energy region, with some parameters behaving

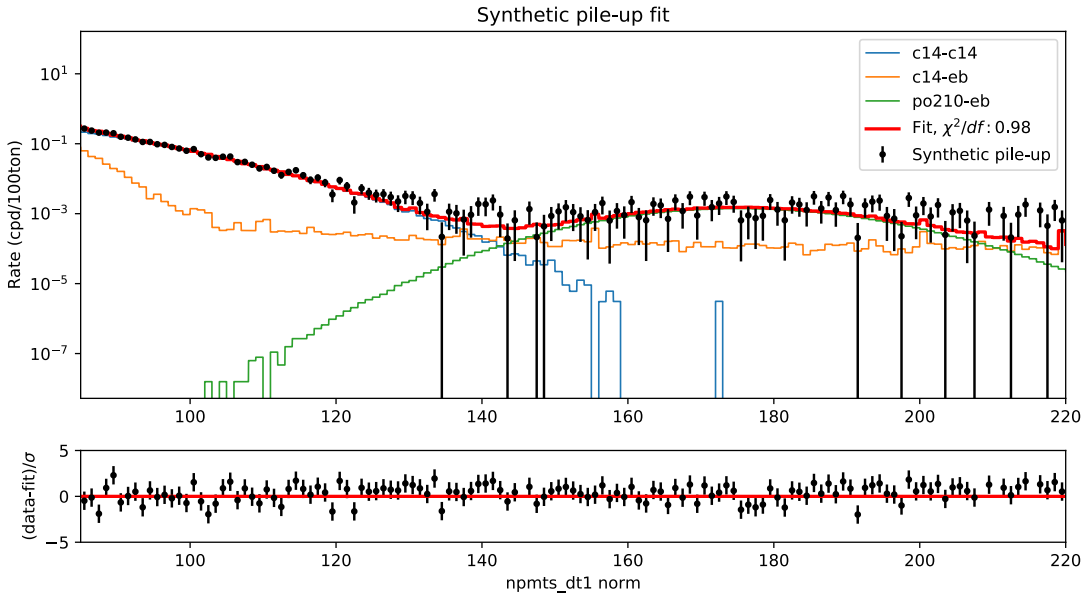


FIGURE 58: The fit of the three different Monte-Carlo pile-up components to the synthetic pile-up spectrum with an E_{min} value of 6. „EB" stands for *external background*. As one can see, the c14-c14 component is, as expected, dominant in the low energies. While c14-EB becomes flat at higher energies, following the energy distribution of single external background events, po210-EB acts as a broader ^{210}Po peak at around 180 *npmts_dt1*. This can be used to calculate the ratios between the rates of the different component, and provide a cross-check to the rate calculation, albeit with less precision compared to the direct calculation from the ^{14}C rate (see text).

erratically under small changes in the fit conditions. It was therefore decided to drop the ^{14}C and ^{14}C-^{14}C pile-up energy range, at the cost of possible sensitivity, albeit gaining confidence in the results, at least until the low-energy region is better understood.

This exclusion has further consequences on the analytical fit since it is very sensitive to the valley between the ^{14}C and ^{210}Po peaks which ends up acting as a discriminator for the rate of ^{85}Kr, one of the most important backgrounds in that energy region. Additionally, while in the Monte-Carlo fit parameters like the energy scale and light yield are tightly constrained to values from calibration data, in the analytical approach these parameters are left free. As a consequence, the sensitivity to ^{85}Kr influences the determination of the energy scale in the analytical fit, and thus the overall count rate in the low-energy region [181]. On that account, using these settings, *i.e.* fitting in an energy range starting at 140 *npmts_dt1*, the analytical fit could not be used anymore. In other words, the CNO analysis (see section 3.7), in Phase III, was made exclusively using the Monte-Carlo fit approach.

B The discovery of CNO neutrinos

B.1 3D corner plot

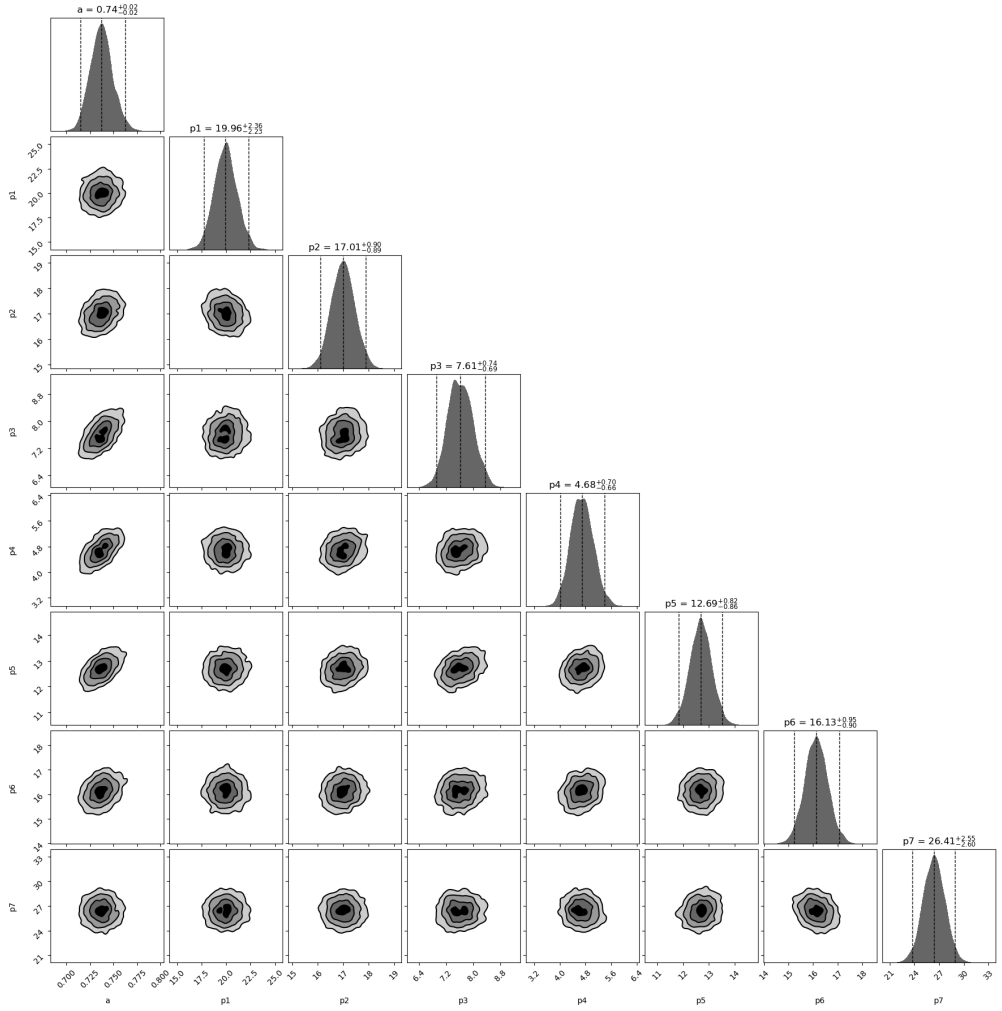


FIGURE 59: Results of a high-dimensional spline fit. The contours were plotted using MultiNest software[229–231]. The left-most parameter a corresponds to the parabolic shape parameter along the ρ^2-axis (See Eq. 30), and all others are the height parameters of the knots, proportional to the rate of ^{210}Po decays. As one can see, all parameters are nicely found and Gaussian confidence intervals can safely be quoted. In general, no spline parameters are correlated, except for some light correlation with a.

B.2 Best-fit curve

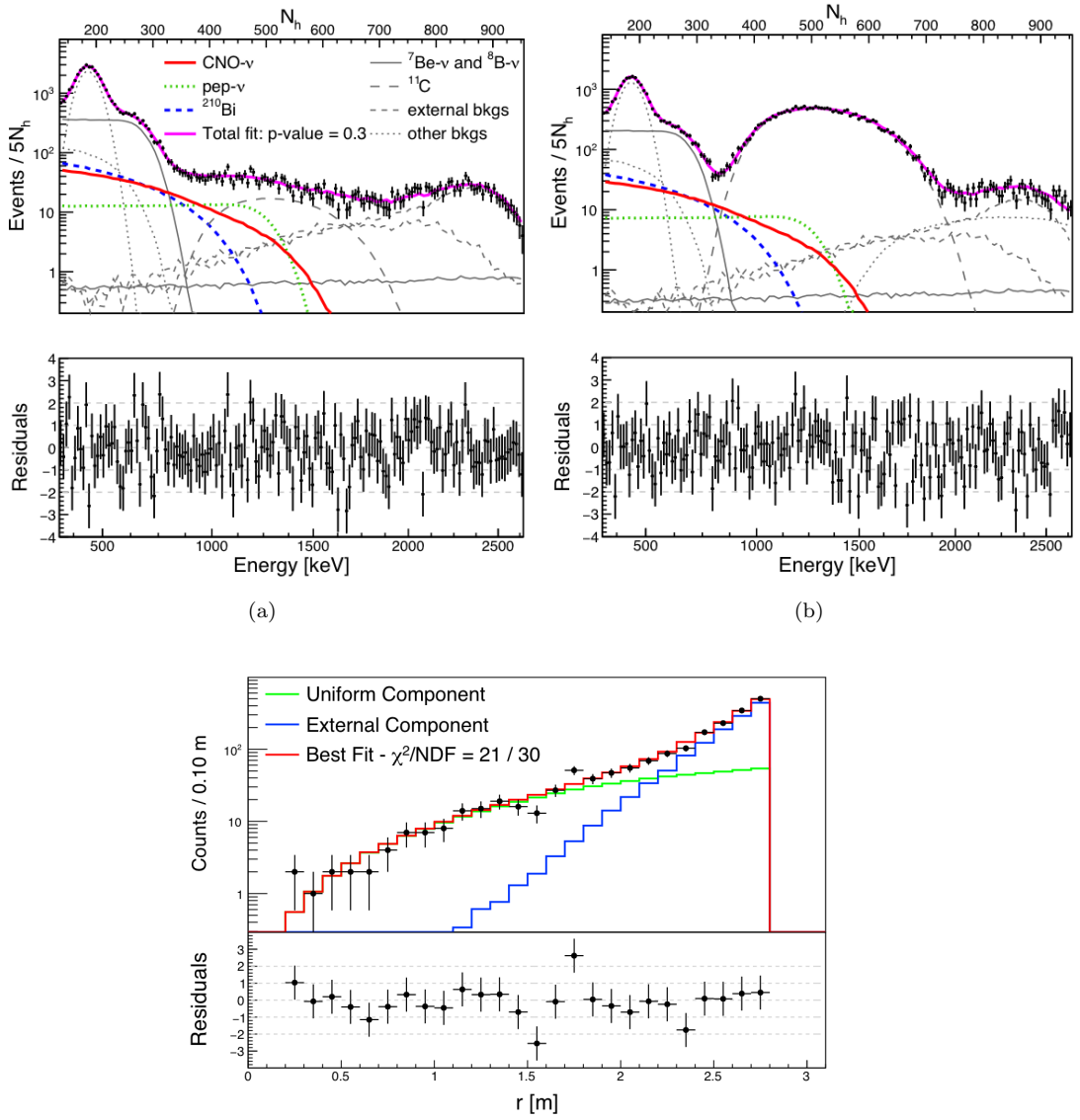


FIGURE 60: The results of a multivariate fit to the Borexino Phase-III data. (a) shows the fit to the TFC-subtracted histogram, depleted in ^{11}C. The red line corresponds to the best-fit to CNO, the dashed blue line to ^{210}Bi, the dotted green line to pep neutrinos, the black points to data and the purple line to the total fitted spectrum. Finally the grey curves show different kinds of backgrounds. (b) shows best-fit to the TFC-tagged spectrum, rich in ^{11}C. (c) shows the results of the radial fit, where one can see that it nicely constrains the rate of externals compared to the rate of singles in the fiducial volume.

C JUNO: solar neutrino correlation plot

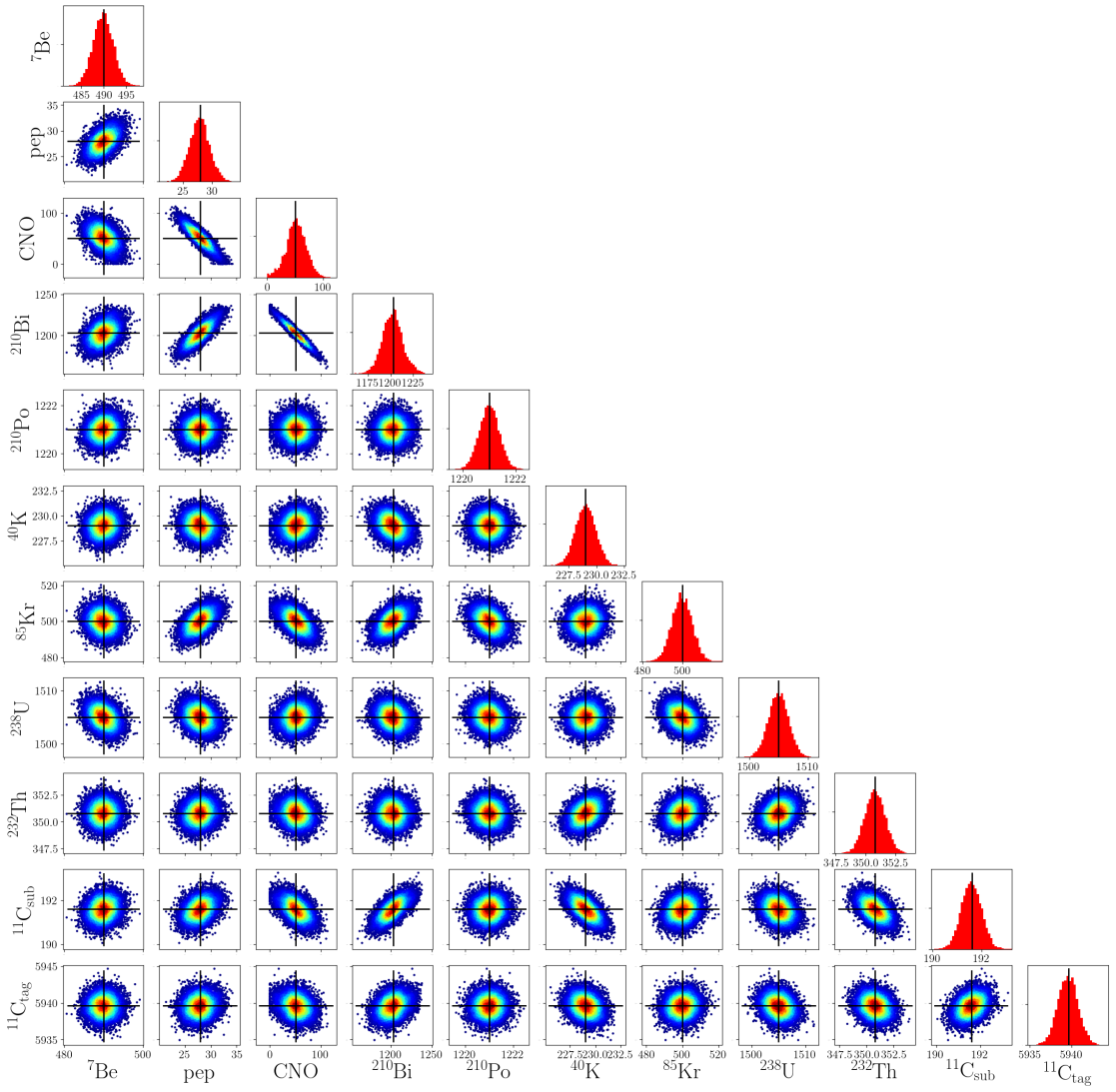


FIGURE 61: Analysis result (see section 4.4.1) for 10.000 fits in the *baseline* scenario for six years of data-taking. All values are in cpd/kton. The black lines indicate the values that were injected in the creation of the datasets. As one can see, in this scenario all components are reconstructed with no bias. Some correlations can be easily recognised.

D OSIRIS Calibration: LED measurements

D.1 Angular profile

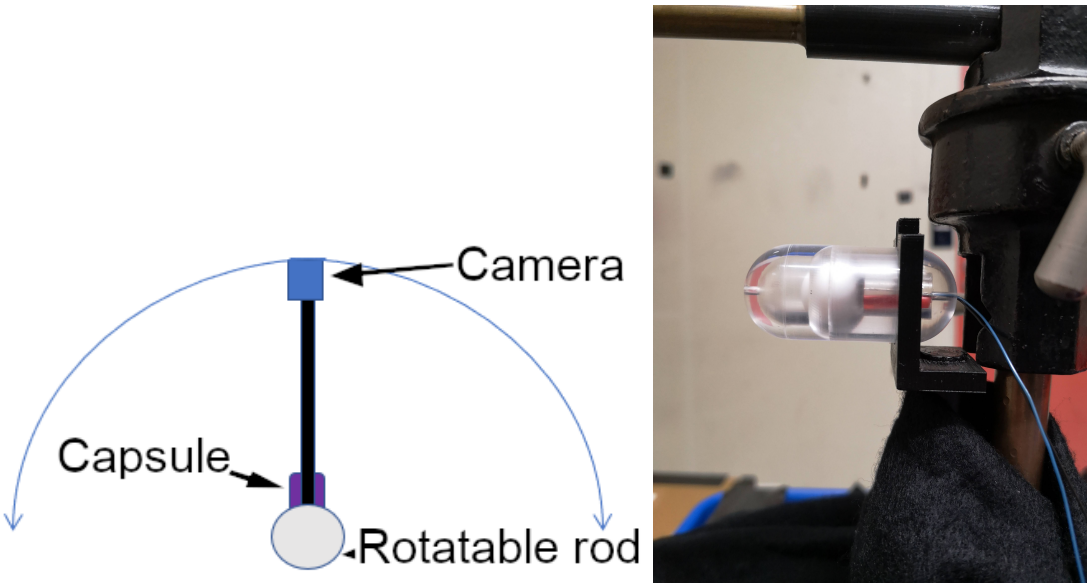
This section describes a measurement series that was performed by Tobias Sterr at the Eberhard Karl's Universität Tübingen, and the subsequent data analysis by myself. The goal of the measurement is to get a quantitative grip on what the angular distribution of the LED-capsule setup is. It is in general advantageous for the light emitted by the capsule to be as isotropic as possible, but practically this is not always possible. Since the subsequent LED-related calibration studies are all affected by this distribution, I organised a measurement.

Experimental setup The LED capsule, see 40, is symmetrical along ϕ (assuming a spherical coordinate system where the z axis goes through the centre of the longest side). Given that the light output from the diffuser is near-perfectly isotropic, the main source of anisotropies is along θ and stems from the geometry of the capsule. Some effort was put to minimise these effects, by e.g. grinding the acrylic surfaces of the capsule to maximise transparency, and by designing the capsule to be as thin as is practicable. However, since some cables have to be inserted inside of the capsule, and not less a LED, perfect isotropy is not achievable. The goal of this measurement is thus to determine achieved amount of isotropy by measuring the angular profile along the θ direction.

The LED capsule was inserted in an custom-designed 3D printed black holder, which was itself supported by a metallic arm to firmly hold the capsule in place. A CCD camera facing the capsule was placed at the end of another arm, which could be rotated 180° around the capsule. Three measurement series were then performed, each with the capsule-holder setup in a different orientation, to minimise shadowing of the capsule through the holder. The measurement series were centered at the front, side, and back of the capsule, respectively, with the front being defined as the direction in which the LED is pointing. The entire setup was placed in a room with good light isolation. A sketch and parts of the assembly can be seen in fig. 62.

Data analysis The measurement resulting in 38 pictures per measurement series: one picture with no light on, as a background measurement, and one picture with the LED as the sole light source for each angle. In each series, the 180° range was sampled in 5° increments. In the analysis, each picture then underwent the same treatment:

1. Subtract the background image from the current image.



(a) (b)

FIGURE 62: (a) A sketch of the measurement setup in Tübingen, as seen from the top. (b) A picture of the setup, zoomed in on the capsule inside of the holder. The rotatable rod is visible on the right side of the capsule, as is the rod holding the camera above.

2. Get all pixel values and insert them in a histogram. A pixel value is basically the amount of light seen in that pixel.

3. Fit a Gaussian to the pixel values. This is necessary since just taking the mean/standard deviation of all values is more easily affected by the values from the capsule - which do not follow the background distribution. This allows to disentangle and results in two values: μ_{bkg} and σ_{bkg}, representing the background noise.

4. Apply a Gaussian filter to the image (in python - skimage.feature.canny). The purpose of a Gaussian filter algorithm is to detect edges in a given image. The canny algorithm employed in this analysis first smooths neighbouring pixel values by applying a Gaussian filter, then calculates gradients in each direction (x and y). Using the gradients the algorithm then looks for local maxima to classify pixels as belonging to so-called "edges". As a final step, the algorithm removes all isolated high-gradient pixels, thus keeping only lines belonging to actual edges in the picture, and removes classified "edges" based on a cut on the minimum value of the gradients belonging to the edge. Should an edge not have any gradients above a user-defined value, the edges are removed from the selection.

5. Convert the output edges to 2D points in the image.

6. Apply a convex hull algorithm to the points. A convex hull algorithm is an algorithm that finds the smallest convex shape capable of enclosing all points in a

given set of points. A famous if outdated example of a convex hull algorithm is the gift wrapping algorithm [261]. A visual example of all the aforementioned steps can be seen in fig. 63.

7. For each pixel, determine if it is inside of the convex hull, then count the number of pixels in the area and take the integral *i.e.* the sum of those pixel values, this results in two values: N_{pixel} and Γ.

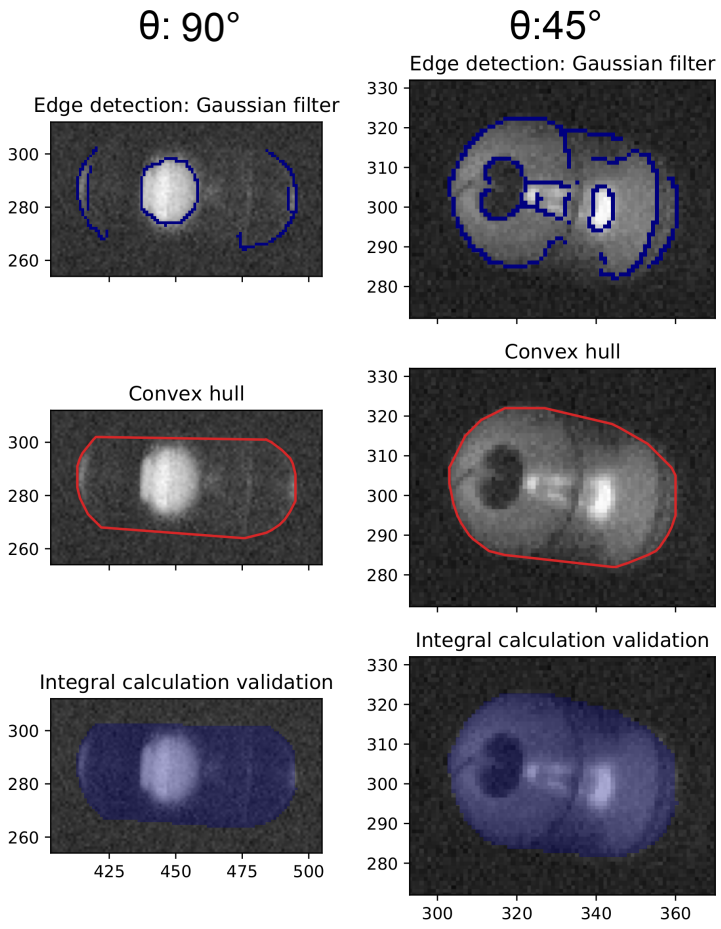


FIGURE 63: Step-by-step of the analysis of the Tübingen data for two different measurements at θ =90° (left) and 45° (right). The first step shows the "edges" output of the Gaussian filter (see text). The second step shows the result of the convex hull algorithm, and the last step shows which pixels were used in the calculation of the integral.

Using the results from the previous steps, one can now make the simple assumption that the pixels used in the integral calculation contain two components: signal + background. The signal should be proportional to the amount of light emitted by the capsule in the corresponding direction, and the background should be the same everywhere. This leads

to the following equation:

$$\Gamma = \alpha + \mu_{bkg} \cdot N_{pixel} \tag{57}$$

$$\leftrightarrow \alpha = \Gamma - \mu_{bkg} \cdot N_{pixel}, \tag{58}$$

where Γ is the value of the sum over all pixels in the signal area, which has a total of N_{pixel} pixels, μ_{bkg} is the mean expected background per pixel, and α is proportional to the emitted amount of light from the LED. Using the width of the background distribution (σ_{bkg}, see step 3 above) as an uncertainty on the value of each pixel, it follows that the uncertainty on α is:

$$\begin{aligned} \sigma_\alpha &= \sqrt{\sigma_{bkg}^2 N_{pixel} + \sigma_{bkg}^2 \frac{N_{pixel}^2}{N_{bkg}}} \\ &= \sigma_{bkg} \cdot \sqrt{N_{pixel} \cdot \left(1 + \frac{N_{pixel}}{N_{bkg}}\right)}, \end{aligned} \tag{59}$$

where σ_α is the uncertainty on α and N_{bkg} is the number of pixels used in the determination of μ_{bkg}.

Results When applied to each image and each measurement series, the equations described above result in fig. 64. As one can see, the measurements agree very well with each other in the range between $-130°$ to $0°$. The mismatches near $-160°$ and $90°$ stem from some light being blocked by the holder. Overall, a difference of a factor of about three was observed between the minimum and maximum α. The front part of the capsule, meaning the range between $-90°$ to $90°$, shows a very homogeneous distribution. The light emission then quickly rises to attain a double peak structure near $-180°$. The double peak structure can be understood as a single peak, originating from the back-side of the diffuser when light from the LED has less if any material to traverse inside of the diffuser, and a dip caused by the cables/legs of the LED itself. A factor three is more significant than anticipated, which is why it is taken into account in the simulation. A more elaborate discussion on the consequences of anisotropy can be found in section 5.4.

D.2 Time and charge behaviour

In order to perform the LED calibration in OSIRIS, one has to be able to control the amplitude of the LED on-the-fly (see section 5.4). As explained in section 5.2.3, this is done by varying a negative DC voltage that is given to a special driver board. The exact value of this voltage is unknown, as is the voltage dependence on the mean number of emitted photons per pulse. In order to determine this and, at the same time, to perform

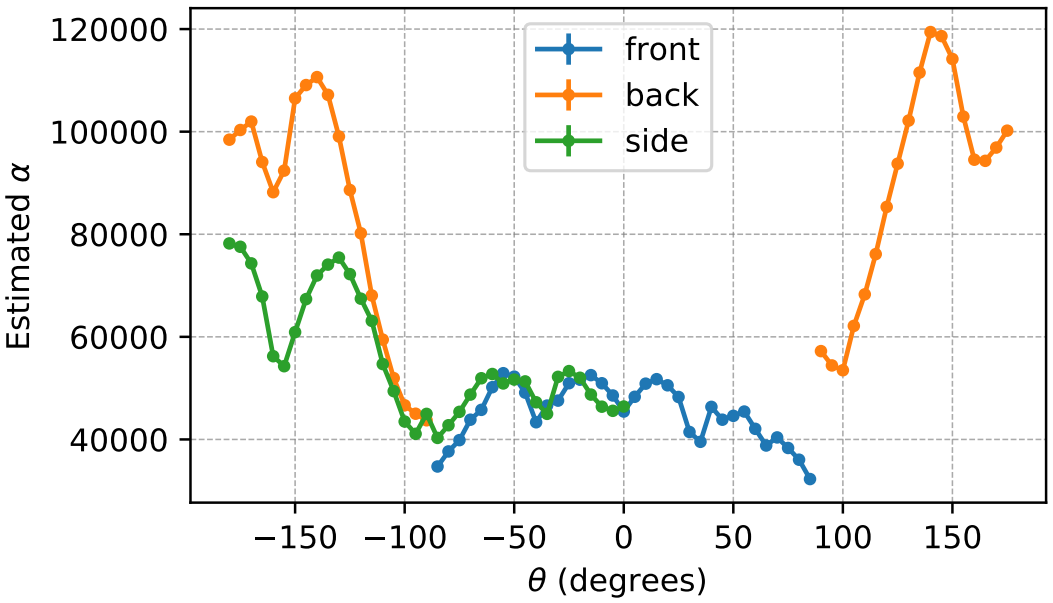


FIGURE 64: Results for α (proportional to amount of light in each image) of the three measurement series. 0° corresponds to looking at the capsule from the front *i.e.* looking at the LED head-on. As one can see, the measurements agree very well with each other in the range between −130° to 0°. The mismatches near −160° and 90° stem from some light being blocked by the holder.

a feasibility test of the LED setup and of the driver board, a series of measurements were taken with a PMT setup in the Physikalisches Institut 3.B. of the RWTH Aachen, thanks to the help of Christian Wysotzki and Jochen Steinmann. This section describes this experiment, the methods used to analyse the data, their results, and concludes with a discussion about LED timing and stability in OSIRIS.

Experimental setup In order to test the LED, a relatively simple setup was used whereby an external trigger (TTL pulse) was laid out to trigger the LED driver board and the readout electronics of a prototype iPMT. Through the VULCAN chip in the iPMT electronics, 500 Mio. samples can be read per second, one sample every two nanoseconds. An adjustable DC power source was placed close to the readout computer to give the ability to quickly change the amplitude of the LED or to turn it off if needed. The LED was placed in front of a PMT which was itself in a well light-isolated room: in this case a container with mostly black surfaces inside - with small feedthroughs to allow for cables to run through the container wall without any light leak being created in the process. The PMT used in the experiment was a 20″ Hamamatsu PMT, identical to the ones used in OSIRIS. A sketch of the setup can be seen in fig. 65. During the measurement process, the LED had to be moved sometimes. Special care was of course taken as to not damage the PMT - using red LEDs as the only source of light inside of

the container and turning off the PMT high voltage whenever the door to the container had to be opened. To avoid measuring all too much dark noise the PMT was given a few minutes to "rest" after each turning-on of the high voltage. This still results in a lot of dark noise - PMTs usually need to "cool off" for hours - but since the LED-based setup allowed us to seek event in clearly defined and narrow time windows, it was not a problem.

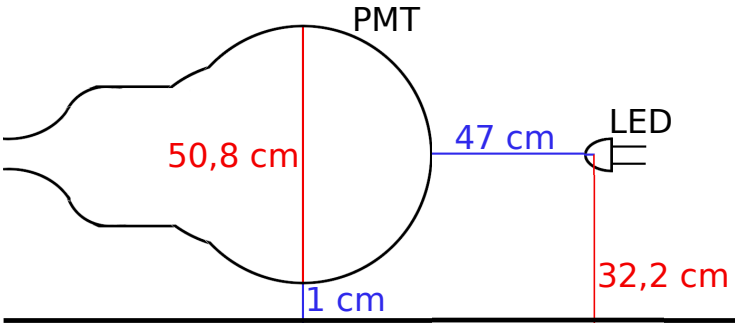


FIGURE 65: A sketch of the PMT-LED setup used in Aachen. The drawing is not to-scale.

Data analysis Several measurements were taken using the PMT-LED setup, with varying voltages set to the LED driver. Specifically, the voltage was decreased from $-4.2\,\mathrm{V}$ to $-4.48\,\mathrm{V}$ in increments of $0.02\,\mathrm{V}$. Each measurement contains at least 100.000 LED triggers - and each trigger contains a waveform with 120 bins (of 2 ns width) containing the ADC output from the VULCAN chip, totalling in a 240 ns measurement window, as well as some metadata surrounding the measurement *e.g.* the UNIX timestamp (with nanosecond precision) of each trigger. During these measurements, the LED driver was pulsed with a frequency of 1 kHz.

The first goal of this analysis was to find a way to detect peaks in the waveforms, see *e.g.* fig. 66, and to assign a timestamp to each detected peak. Peak finding was done relatively simply by looking at the expected deviation from the baseline in each time bin: for each measurement series, 10.000 waveforms were first projected onto the y-axis and a Gaussian was fitted to extract the position of the baseline. This resulted in a position and standard deviation for the baseline. In order to find peaks, the idea was that in each waveform, the ADC bins could now be converted to a distance from the baseline in units of the standard deviation. By setting cuts on said deviation, one can find regions of so-called "high-σ". To elaborate, a "high-σ" region is opened when a bin is over the threshold, and is closed again when a bin goes under one standard deviation. This ensures that peaks are identified as one single entity. Of course, all this begs the question of where to set the threshold. This question is answered in fig. 67(a). The idea was that, for each measurement series, the total number of detected peaks was

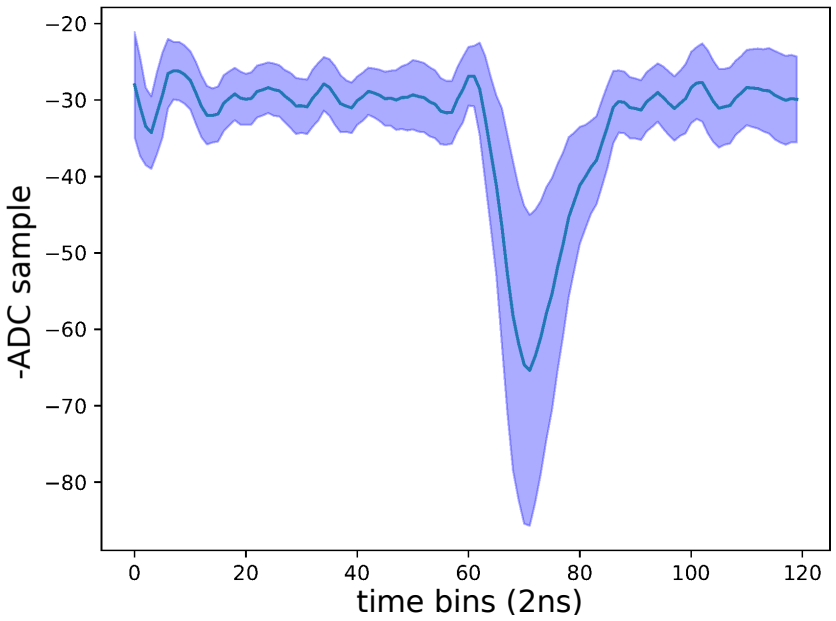
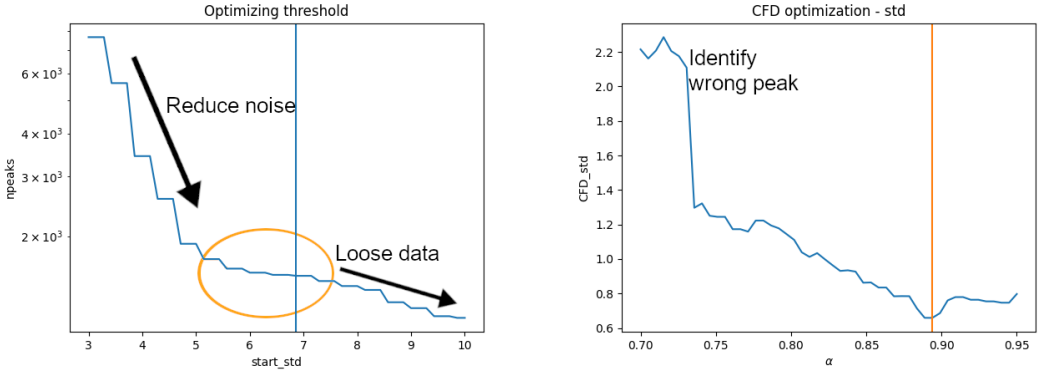


FIGURE 66: Average of all waveforms that contain exactly one hit, taken from a measurement with V=−4.30 V. The shaded area shows the $\pm 1\sigma$ region around the mean. Fluctuations in the baseline can be seen, as can the peak very clearly.

plotted as a function of the threshold. One can observe a fast drop in this number at first, indicating that noise is being adequately rejected. This drop transitions into a plateau region, indicated by the orange oval in the figure, where the threshold is close to optimum. Then, for very high values of the threshold, the number of identified peaks starts to drop again, indicating that it is now in a region where data starts to be lost. I chose the optimal value of the threshold as the point in the curve with the smallest gradient, as indicated by the vertical line in the figure.

Secondly, for each identified peak, a timestamp needed to be extracted. There are many different methods to do this, but I chose to use a simple constant fraction discriminator as it is robust and very effective considering its simplicity. The idea of a constant fraction discriminator is to define the position of the timestamp as the position where the peak reaches a certain percentage of its maximum height. As such, even for different amplitudes, as long as the peaks have similar shapes (which is expected in PMT outputs), the timestamp extraction stays very consistent. Figure 67(b) shows an example of an optimisation of the CFD. As with the previous threshold optimisation, this operation was of course performed for each measurement series. The idea here was to minimise the standard deviation of the extracted timestamps with respect to the CFD percentage value, hereinafter referred to as α, in order to find its optimal value.

Using the information gathered with the aforementioned methods, one can start to produce plots such as the one visible in fig. 68. As one can see, the timestamp distributions



(a) (b)

FIGURE 67: (a) Example threshold optimisation. The blue line shows the number of identified peaks as a function of the standard deviation starting from which a region is identified as a peak. As one can see, the number of identifications drops significantly up to a deviation of about five, then flattens out for a bit before starting to drop again after about seven. The optimised threshold value is shown by the vertical line. (b) Example CFD optimisation. The standard deviation of the reconstructed hit-times as a function of the CFD fraction. As one can see, when α is too low, some features in the background are marked instead of the peak, then the standard deviation goes down to a minimum at about $\alpha = 0.9$. The optimal value is shown by the vertical orange line.

agree quite well with expectations: a Gaussian from the LED arrival times plus a constant from the dark noise. This very powerful fit allowed us to distinguish between peaks belonging to the LED signal and peaks belonging to the background. It is important to note that the number of peaks extracted from this fit is in general not equal to the number of photo-electrons that were recorded by the iPMT. Indeed, several photons arriving at (nearly) the same time on the cathode would result in a single peak - and up to now I was simply counting the number of peaks. However, it is possible to use the nature of Poisson statistics to my advantage: the number of waveforms with no signal is also an information about the signal rate! Let μ^S be the signal-occupancy on the PMT *i.e.* the average number of signal-photons per waveform. It is important to note that μ^S is not only determined by the amplitude of the LED pulses and the geometry of the setup, but also by the quantum efficiency of the PMT. Then, the probability of seeing no signal peak in a given waveform is given by:

$$P_{\mu^S}(X = 0) = e^{-\mu^S}. \tag{60}$$

This formula can be re-writen as:

$$\mu^S = -ln\left(\frac{N - N_S}{N}\right), \tag{61}$$

where N is the total number of recorded waveforms and N_S is the reconstructed number of signal peaks. I observed that some waveforms could have multiple peaks - however

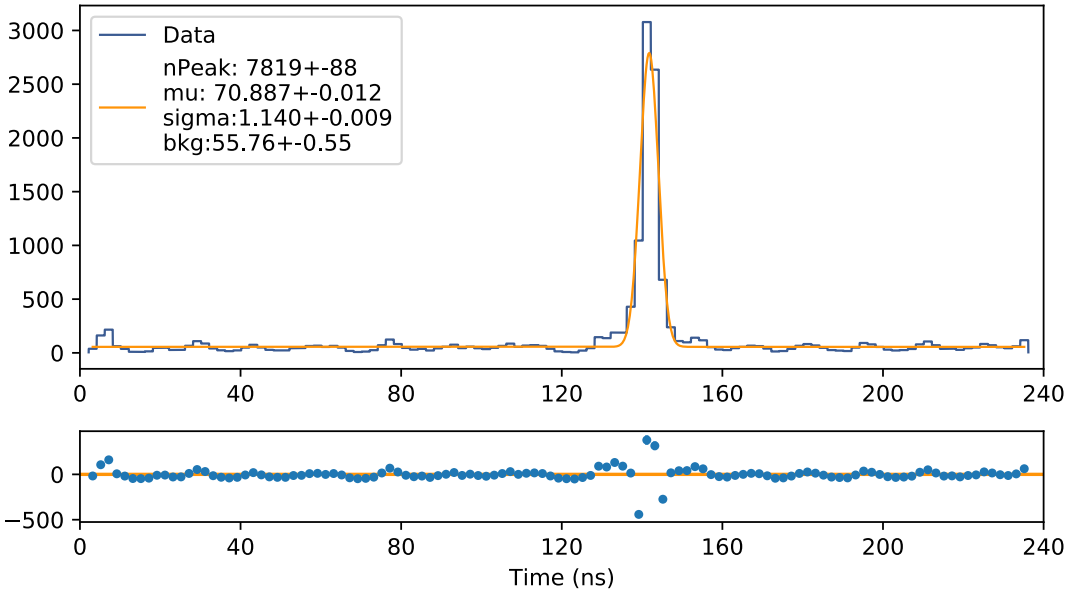


FIGURE 68: Histogram of all the timestamps recorded from a measurement series taken with a voltage of −4.3 V. The distribution was fitted with the sum of a Gaussian and a constant, since this is what would be expected from LED arrival times and dark noise, respectively. The fitted parameters are, in that order, the number of events under the Gaussian curve, the position and width of the Gaussian, and the mean expected hits per bin from background. One should keep in mind that while the x-axis is in nanoseconds, the fit was performed in units of the VULCAN samples which result in 2 ns bins.

given the proximity of the signal peaks (within a few nanoseconds) - there is no chance for the peak-finding algorithm described above to recognise them as separate. It is therefore safe to consider the reconstructed number of signal peaks as equal to the number of waveforms with at least one hit. For completeness, the uncertainty on the μ^S from eq. (61) is given by:

$$\sigma_{\mu^S} = \frac{\sigma_{N_S}}{\sqrt{N - N_S}}, \tag{62}$$

where σ_{N_S} is the uncertainty, given by the fit, on N_S. Given the relatively high statistics, this number is of course very close to $\sqrt{N_S}$, but somewhat different because of the effect of the shape information.

Another calculation that can be made is to calculate the integral of every waveform. This allows one to not only cross-check the previous result on μ^S, but also to verify ("sanity-check") the validity of the data from the PMT. In order to get this integral, the waveforms were simply integrated over their entirety and the results inserted in a histogram. I used a procedure based on [251] that does not simply fit one or two Gaussians in intuitive areas, as is often seen done on charge histograms, but actually incorporates knowledge about the physical processes inside of the PMT that allows to directly extract information about not only the measured occupancy, but also some of the properties of the PMT. However, different approximations are possible or even

necessary depending on the situation. For simplicity, the following paragraph describes the fit procedure that was used without going into a detailed discussion of the differences with respect to [251].

The fit does not take discrete processes, such as thermo-emission, into account, for the simple reason that for the iPMT these effects seem to be low enough that I could not find any sensitivity to them. With those good news in mind, the fit function simply becomes a weighted sum of Gaussians. In principle, when no photons are detected, one would solely observe the so-called "Pedestal" which has a non-zero width due to low charge processes and electronic noise which are necessarily present in every event [251]. Let the pedestal be described by a Gaussian of shape: (q_0, σ_0). Both of these parameters are dependent on properties of the PMT including gain. A peak containing a single photon should then in principle incite an additional charge q_1 distributed with a Gaussian of width σ_1. For a peak containing more than one photon, this charge distribution would simply be sampled from more often. It is at this point that the knowledge of Poisson statistics becomes useful once again: for a given occupancy of μ (μ is not only the probability of observing a signal event but also a background event, thus $\mu = \mu^S + \mu^B$, where μ^B is the expected number of dark-noise-induced peaks per waveform), the height of the Gaussians will be distributed according to the Poisson probabilities. This can be summarised as following:

$$f(x) = \sum_{n=0}^{\infty} P_\mu(X = n) \cdot G(x; q_0 + nq_1, \sqrt{\sigma_0^2 + n\sigma_1^2}) \tag{63}$$

where $P_\mu(X = n)$ is the Poisson probability with an expected value of μ to observe n, and $G(x, \mu, \sigma)$ is a Gaussian probability density function with an expected value of μ and a standard deviation of σ. Of course, for practical reasons, this sum can not be taken up to infinity. Instead, I developed an algorithmic procedure whereby the data is first fitted using $n = 2$, after which n is iteratively incremented by one in each step. This is repeated until the χ^2/dof value from the fit in one iteration, say iteration n_{max}, is higher than that of the previous iteration. The results of the fit with $n_{max} - 1$ were then used as a final value. An example of such a fit can be seen in fig. 69.

As one can see, the fit is a very good match to the data. Similar agreement was observed in all measurement series. This gives great confidence in the output from the PMT - not only that the PMT behaves appropriately with no hiccups, but also that our understanding of the physics of the PMT is adequate. This also allows to reconstruct μ with a high precision using the entirety of the dataset. This brings us to the cross-check of μ^S. On its own, the charge fit cannot differentiate between signal and dark noise events. It is however possible to combine the information from the charge fit and from the time fit (see fig. 68). One can simply combine the fact that $\mu = \mu^S + \mu^B$, and that

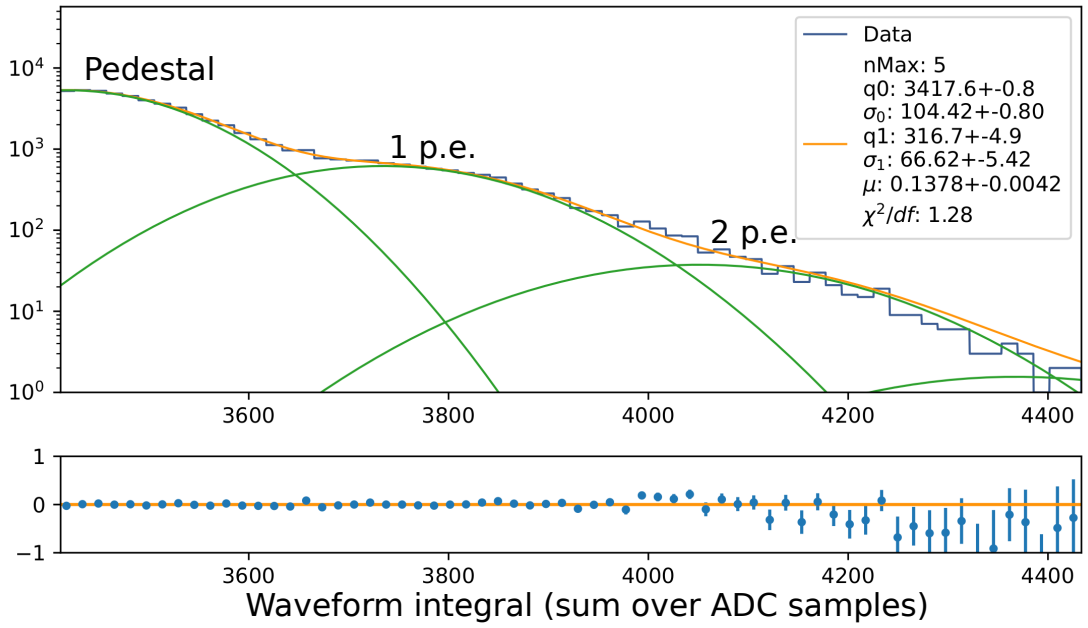


FIGURE 69: An example charge fit on a measurement taken with a voltage of 4.3 V. As one can see the fit (eq. (63)) is a good match to the data, and the value of μ (see text) could be inferred with a high precision of just about 3%. This result really underlines that the data coming from the PMT is well understood - and of course that the PMT itself behaves as expected.

similarly to what was seen with μ^S:

$$P_{\mu^B}(X = 0) = e^{-\mu^B}. \tag{64}$$

This directly leads to:

$$\mu^S = \mu + ln\left(\frac{N - N_B}{N}\right), \tag{65}$$

where N_B is the number of identified background peaks. According to the values of μ^B that were determined (using a formula analogous to eq. (61)), the probability of two-or-more dark noise hits occurring in a single waveform is between 0.2% and 0.5%. This does not include the probability that the peaks would most often be far enough apart to be recognised as separate. It is therefore a good approximation to consider N_B as the number of dark noise hits as well, making eq. (65) valid. For completeness, the following formula gives the uncertainty on the μ^S from eq. (65):

$$\sigma_{\mu^S} = \sqrt{\sigma_\mu^2 + \frac{\sigma_{N_B}^2}{N - N_b}}. \tag{66}$$

The determined values of μ, μ^S and μ^B can be found in Table 20.

Voltage (V)	$\mu \cdot 10^1$	$\mu^S_{time} \cdot 10^2$	$\mu^B \cdot 10^2$	$\mu^S_{charge} \cdot 10^2$
-4.20	0.668 ± 0.037	0.774 ± 0.040	6.57 ± 0.12	0.705 ± 0.038
-4.30	1.371 ± 0.015	8.252 ± 0.032	6.545 ± 0.028	7.643 ± 0.089
-4.32	2.095 ± 0.055	13.84 ± 0.13	6.261 ± 0.083	14.43 ± 0.41
-4.34	2.929 ± 0.062	22.40 ± 0.18	6.887 ± 0.087	22.40 ± 0.54
-4.35	3.57 ± 0.14	23.15 ± 0.26	8.56 ± 0.14	26.1 ± 1.1
-4.36	4.023 ± 0.067	33.36 ± 0.24	7.431 ± 0.091	32.89 ± 0.65
-4.38	6.35 ± 0.13	52.05 ± 0.34	8.74 ± 0.10	54.4 ± 1.3
-4.40	9.51 ± 0.20	85.78 ± 0.57	10.32 ± 0.11	84.8 ± 2.0

TABLE 20: Determined values of μ, μ^S and μ^B for different voltages. μ is a direct result from the charge fit, μ^S_{time} and μ^B were calculated using the "time-fit" results with eq. (61). μ^S_{charge}, shown here as a cross-check of both methods, was calculated by combining charge-fit and time-fit results with eq. (65). The values are quite compatible.

Results The results of the previous section were used to estimate the number of emitted photons per LED pulse. This calculation necessitates some assumptions. The first assumption is perfect isotropy of the light emission from the capsule. The second is that the detection efficiency is identical everywhere on the PMT cathode, and that the quantum efficiency is exactly 0.3 (though this is based on existing measurements). A Toy-MC reflecting those assumptions was used to calculate the factor between the number of emitted photons from the capsule and the number of detected photo-electrons. This results in fig. 70. As one can see, the results from the μ^S calculation from eq. (61) and eq. (65) agree very well with each other. Secondly, there is a clear logarithmic relationship between the number of emitted photons per pulse and the voltage on the LED driver. This is as expected and as one can see, this allows us to control the amplitude of the LED over several orders of magnitude. As will be shown in later sections, the number of photons that the LED must emit per pulse is, depending on the situation, between about 10 and 50, which is well in the investigated range. Taking into account the (advertised) sensitivity of the card that will be used to generate the DC power (NI PCI 6703[28]), the gradient of this curve near $n = 50$ results in an uncertainty of ± 1 photon per year of measurement. From the aforementioned assumptions it is clear that comparatively large systematics shifts are expected to be present in from fig. 70, which shows only statistical uncertainties. Systematics include peak detection efficiency and assumptions on the Toy-MC. This is no problem here because the main statement of the plot would be unaffected. Further, the curve is not intended to be used as a reference during the experiment (although it is an important proof-of-concept for the setup), except to get an intuition of the expected numbers and sensitivity beforehand.

The reason such a reference cannot be created in advance is two-fold: 1. The conditions inside of the AV of OSIRIS are not really reproducible in a lab. 2. It is known from the experience of Daya-Bay scientists that the amplitude of the LED pulses depends

[28] https://www.ni.com/de-de/support/model.pci-6703.html

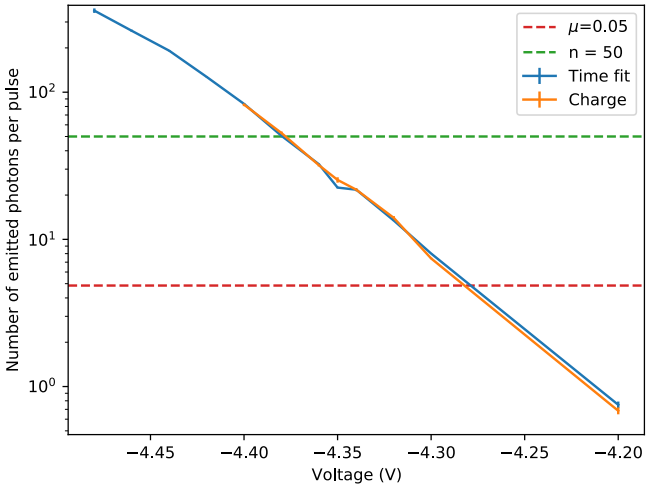


FIGURE 70: Number of emitted photons from the LED capsule per pulse as a function of the voltage on the driver. The blue and orange lines show the results using μ^S from eq. (61) and eq. (65), respectively. The red dotted line corresponds to a signal occupancy of $\mu^S = 0.05$ on the PMT used in the setup in Aachen. The green dotted line show where n=50 photons are emitted per pulse. As comparison, in OSIRIS between a 10 and 50 photons will be emitted per pulse, depending on the position of the capsule, to reach an occupancy of $\mu^S = 0.01$ or less on the PMTs.

on the temperature of the LS (which is z-dependent in OSIRIS) and on how much of the cable is still on the cable spool VS how much is deployed. However, this can easily be adjusted on-the-fly during a calibration run. Since our online DAQ software is able to give us near-instantaneous information, one can quickly get the occupancy on each PMT while a calibration run is ongoing. The voltage to the LED can then be slowly and gradually increased until a signal can be seen. From these measurement one could infer that good values would be $-4.2\,\mathrm{V}$ for the starting voltage and $0.01\,\mathrm{V}$ to $0.02\,\mathrm{V}$ for the step size. It is important to start with as low a voltage as possible (absolute value), since a very high voltage could in worst cases damage the PMTs (there is no danger to the LED itself, as the driver board was thoroughly tested up to almost $-50\,\mathrm{V}$, and the DC generator used in OSIRIS is only able to deliver currents down to $-10\,\mathrm{V}$).

Timing consideration Another effect that can be measured with this setup is whether one can observe a time dependence on the angle of exit from the capsule. Overall, given the length of the capsule (about 4 cm) and the refraction index of acrylic (about 1.5), one can expect some $\mathcal{O}(0.1\,\mathrm{ns})$ differences as a function of the angle, which is no problem given the TTS requirement (see section 5.1). Some problems could however arise should a significant portion of the emitted light get "trapped" in the capsule, meaning if there is one angle at which the light can reflect several times inside of the acrylic before leaving the capsule. In the images taken in the Tuebingen series of measurements (see

Appendix D.1), we (Tobias Sterr and myself) could very clearly see the outlines of the diffuser from almost all angles, indicating that the light takes a direct path out except in some positions near 50° where some fuzziness was observed (the angle is relative to the front of the capsule *i.e.* 0° would be looking directly opposite to the direction in which the LED is pointing). This probably only stems from light being able to travel more than one path and end up in the same angle outwards, but nevertheless we decided to investigate it further.

In order to measure this possible effect, several more measurements were taken between which the only differences were the orientation of the LED relative to the PMT. The voltage was set to $-4.30\,\mathrm{V}$ as it was previously observed that, at this voltage, the PMT was illuminated with an occupancy of less than 0.05, which guarantees a high-purity sample of single p.e. events. This is a must for a reliable and precise timing measurement [192]. Because it was clear that high statistics were needed in order to perform this measurement and because I had limited time on the day of those measurements, I decided to remove the LED driver from the setup (because we could only use it up to a frequency of 1 kHz - resulting in about 17 minutes of measurement time needed for each angle, plus the time needed to turn off the high voltage, move the LED, and reboot the setup) and use pulses from the pulse generator directly. This was possible because of the capabilities of the pulse generator that was available on-site. After some quick testing, it became clear that the lower precision (the led driver produces, by design, much sharper peaks than the pulse generator) would quickly be compensated for by the high statistics, as we were able to pulse the LED with 10 kHz.

The results of the measurement series can be seen in fig. 71. The results are as expected with a $\mathcal{O}(0.1\,\mathrm{ns})$ difference in time induced by travel time inside of the acrylic capsule and diffuser, and faster times observed from the sides ($\theta = \pm 90°$) where the light can escape more quickly. Unfortunately, the data for other angles was unusable because of a faulty configuration. What one can observe is a $\pm 0.15\,\mathrm{ns}$ time difference between light emitted at different angles, which is well below the TTS requirement.

LED stability The data recorded in this experiment allows to verify one more interesting point for LED measurements and that is the stability of the LED pulses. The intensity of the LED pulses can (and must) be varied on-the-fly during calibration runs, but large variations in the intensity during the runs themselves could be very problematic in keeping consistent occupancies on the PMTs. Since during a weekly calibration a single LED pulsing/measurement is to take less than ten minutes (see section 5.7), the LED must only be stable during those time-scales. The timestamp information from a

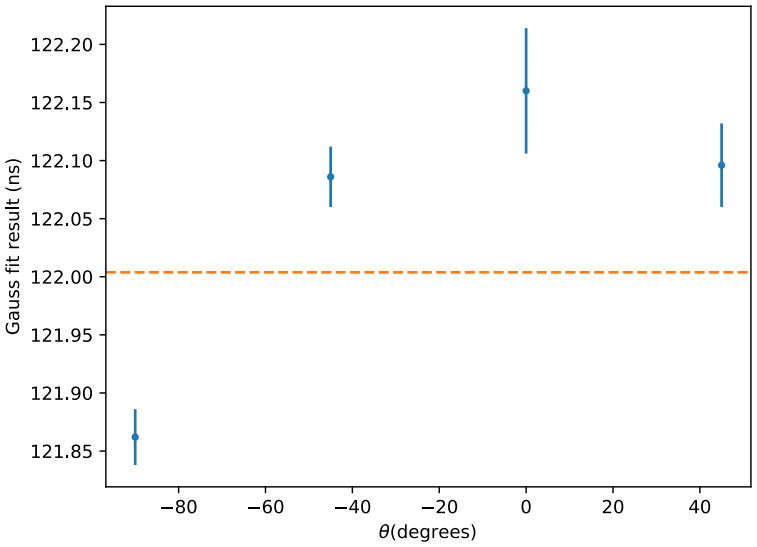


FIGURE 71: Position of the LED peak as a function of the angle. θ =0° corresponds to the emission direction of the LED. The results are as expected with a $\mathcal{O}(0.1\,\text{ns})$ difference in time induced by travel time inside of the acrylic capsule and diffuser, and faster times observed from the sides ($\theta = \pm 90°$) where the light can escape more quickly. The data for other angles was unusable because of a faulty configuration.

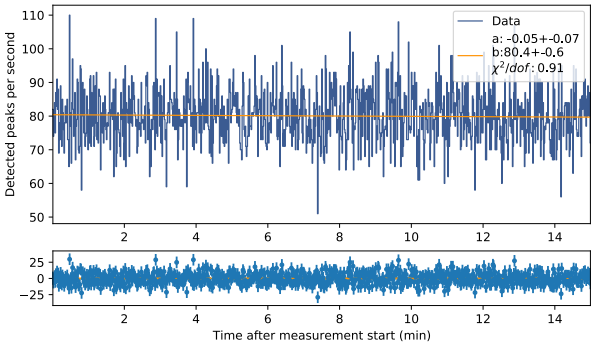


FIGURE 72: Number of detected peaks per second as a function of time. The distribution was fitted with a line ($a \cdot x + b$). From the results it is clear that the measured slope is fully compatible with zero.

large measurement series (900.000 triggers) was used to create fig. 72. As one can see, no instabilities can be seen in a 15 minute time window.

Bibliography

[1] C. D. Ellis W. A. Wooster, The Continuous Spectrum of beta-Rays, *Nature*, 1927 119(2998) 563, http://dx.doi.org/10.1038/119563c0, number: 2998 Publisher: Nature Publishing Group

[2] E. Fermi, Versuch einer Theorie der beta-Strahlen. I, *Z. Physik*, 1934 88(3) 161, http://dx.doi.org/10.1007/BF01351864

[3] M. Aker *et al.*, Direct neutrino-mass measurement with sub-electronvolt sensitivity, *Nat. Phys.*, 2022 18(2) 160, http://dx.doi.org/10.1038/s41567-021-01463-1, number: 2 Publisher: Nature Publishing Group

[4] G. W. Rodeback J. S. Allen, Neutrino Recoils Following the Capture of Orbital Electrons in 37a, *Phys. Rev.*, 1952 86(4) 446, http://dx.doi.org/10.1103/PhysRev.86.446, publisher: American Physical Society

[5] F. Reines C. L. COWANjun., The Neutrino, *Nature*, 1956 178(4531) 446, http://dx.doi.org/10.1038/178446a0, number: 4531 Publisher: Nature Publishing Group

[6] G. Danby *et al.*, Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos, *Phys. Rev. Lett.*, 1962 9(1) 36, http://dx.doi.org/10.1103/PhysRevLett.9.36, publisher: American Physical Society

[7] K. Kodama *et al.*, Observation of tau neutrino interactions, *Physics Letters B*, 2001 504(3) 218, http://dx.doi.org/10.1016/S0370-2693(01)00307-0

[8] Stefan Schael et al., Precision electroweak measurements on the Z resonance, *Physics Reports*, 2006 427(5) 257, http://dx.doi.org/10.1016/j.physrep.2005.12.006

[9] M. Goldhaber, L. Grodzins A. W. Sunyar, Helicity of Neutrinos, *Phys. Rev.*, 1958 109(3) 1015, http://dx.doi.org/10.1103/PhysRev.109.1015, publisher: American Physical Society

[10] B. Pontecorvo, Mesonium and Antimesonium, *Soviet Journal of Experimental and Theoretical Physics*, 1958 6 429, aDS Bibcode: 1958JETP....6..429P

[11] B. Pontecorvo, Neutrino Experiments and the Problem of Conservation of Leptonic Charge, *Soviet Journal of Experimental and Theoretical Physics*, 1968 26 984, aDS Bibcode: 1968JETP...26..984P

[12] Super-Kamiokande Collaboration *et al.*, Evidence for Oscillation of Atmospheric Neutrinos, *Phys. Rev. Lett.*, 1998 81(8) 1562, http://dx.doi.org/10.1103/PhysRevLett.81.1562, publisher: American Physical Society

[13] SNO Collaboration *et al.*, Measurement of the Rate of nu_e+d to p+p+e- Interactions Produced by 8B Solar Neutrinos at the Sudbury Neutrino Observatory, *Phys. Rev. Lett.*, 2001 87(7) 071301, http://dx.doi.org/10.1103/PhysRevLett.87.071301, publisher: American Physical Society

[14] R. Davis, D. S. Harmer K. C. Hoffman, Search for Neutrinos from the Sun, *Phys. Rev. Lett.*, 1968 20(21) 1205, http://dx.doi.org/10.1103/PhysRevLett.20.1205, publisher: American Physical Society

[15] J. N. Abdurashitov *et al.*, Results from SAGE (The Russian-American gallium solar neutrino experiment), *Physics Letters B*, 1994 328(1) 234, http://dx.doi.org/10.1016/0370-2693(94)90454-5

[16] P. Anselmann *et al.*, Solar neutrinos observed by GALLEX at Gran Sasso, *Physics Letters B*, 1992 285(4) 376, http://dx.doi.org/10.1016/0370-2693(92)91521-A

[17] M. Altmann *et al.*, Complete results for five years of GNO solar neutrino observations, *Physics Letters B*, 2005 616(3) 174, http://dx.doi.org/10.1016/j.physletb.2005.04.068

[18] SNO Collaboration *et al.*, Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory, *Phys. Rev. Lett.*, 2002 89(1) 011301, http://dx.doi.org/10.1103/PhysRevLett.89.011301, publisher: American Physical Society

[19] The KamLAND Collaboration *et al.*, Precision Measurement of Neutrino Oscillation Parameters with KamLAND, *Phys. Rev. Lett.*, 2008 100(22) 221803, http://dx.doi.org/10.1103/PhysRevLett.100.221803, publisher: American Physical Society

[20] Particle Data Group *et al.*, Review of Particle Physics, *Progress of Theoretical and Experimental Physics*, 2020 2020(8) 083C01, http://dx.doi.org/10.1093/ptep/ptaa104

[21] Z. Maki, M. Nakagawa S. Sakata, Remarks on the Unified Model of Elementary Particles, *Progress of Theoretical Physics*, 1962 28(5) 870, http://dx.doi.org/10.1143/PTP.28.870

[22] C. Giunti C. W. Kim, Coherence of neutrino oscillations in the wave packet approach, *Phys. Rev. D*, 1998 58(1) 017301, http://dx.doi.org/10.1103/PhysRevD.58.017301, publisher: American Physical Society

[23] Daya Bay Collaboration, Study of the wave packet treatment of neutrino oscillation at Daya Bay, *Eur. Phys. J. C*, 2017 77(9) 606, http://dx.doi.org/10.1140/epjc/s10052-017-4970-y

[24] A. de Gouvêa, V. De Romeri C. A. Ternes, Probing neutrino quantum decoherence at reactor experiments, *J. High Energ. Phys.*, 2020 2020(8) 18, http://dx.doi.org/10.1007/JHEP08(2020)049

[25] J. collaboration *et al.*, Damping signatures at JUNO, a medium-baseline reactor neutrino oscillation experiment, *arXiv:2112.14450 [hep-ex, physics:hep-ph]*, 2021 ArXiv: 2112.14450

[26] S. Parke M. Ross-Lonergan, Unitarity and the three flavor neutrino mixing matrix, *Phys. Rev. D*, 2016 93(11) 113009, http://dx.doi.org/10.1103/PhysRevD.93.113009, publisher: American Physical Society

[27] NuFIT | NuFIT, 2021

[28] I. Esteban *et al.*, The fate of hints: updated global analysis of three-flavor neutrino oscillations, *J. High Energ. Phys.*, 2020 2020(9) 178, http://dx.doi.org/10.1007/JHEP09(2020)178

[29] Super-Kamiokande Collaboration *et al.*, Solar neutrino measurements in Super-Kamiokande-IV, *Phys. Rev. D*, 2016 94(5) 052010, http://dx.doi.org/10.1103/PhysRevD.94.052010, publisher: American Physical Society

[30] SNO+ Collaboration *et al.*, Measurement of the 8b solar neutrino flux in sno+ with very low backgrounds, *Phys. Rev. D*, 2019 99(1) 012012, http://dx.doi.org/10.1103/PhysRevD.99.012012, publisher: American Physical Society

[31] KamLAND Collaboration *et al.*, Reactor on-off antineutrino measurement with KamLAND, *Phys. Rev. D*, 2013 88(3) 033001, http://dx.doi.org/10.1103/PhysRevD.88.033001, publisher: American Physical Society

[32] A. Albert *et al.*, Measuring the atmospheric neutrino oscillation parameters and constraining the 3+1 neutrino model with ten years of ANTARES data, *J. High Energ. Phys.*, 2019 2019(6) 113, http://dx.doi.org/10.1007/JHEP06(2019)113

[33] IceCube Collaboration *et al.*, Measurement of Atmospheric Neutrino Oscillations at 6–56 GeV with IceCube DeepCore, *Phys. Rev. Lett.*, 2018 120(7) 071801, http://dx.doi.org/10.1103/PhysRevLett.120.071801, publisher: American Physical Society

[34] Super-Kamiokande Collaboration *et al.*, Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV, *Phys. Rev. D*, 2018 97(7) 072001, http://dx.doi.org/10.1103/PhysRevD.97.072001, publisher: American Physical Society

[35] K. Abe *et al.*, Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations, *Nature*, 2020 580(7803) 339, http://dx.doi.org/10.1038/s41586-020-2177-0, number: 7803 Publisher: Nature Publishing Group

[36] NOvA Collaboration *et al.*, First measurement of neutrino oscillation parameters using neutrinos and antineutrinos by NOvA, *Phys. Rev. Lett.*, 2019 123(15) 151803, http://dx.doi.org/10.1103/PhysRevLett.123.151803, publisher: American Physical Society

[37] P. Adamson *et al.*, Combined analysis of nu¨mu disappearance and nu¨mu to nu¨e appearance in minos using accelerator and atmospheric neutrinos, *Phys. Rev. Lett.*, 2014 112(19) 191801, http://dx.doi.org/10.1103/PhysRevLett.112.191801, publisher: American Physical Society

[38] The Daya Bay Collaboration *et al.*, Measurement of the Electron Antineutrino Oscillation with 1958 Days of Operation at Daya Bay, *Phys. Rev. Lett.*, 2018 121(24) 241805, http://dx.doi.org/10.1103/PhysRevLett.121.241805, publisher: American Physical Society

[39] H. de Kerret *et al.*, Double Chooz theta13 measurement via total neutron capture detection, *Nat. Phys.*, 2020 16(5) 558, http://dx.doi.org/10.1038/s41567-020-0831-y, number: 5 Publisher: Nature Publishing Group

[40] C. D. Shin *et al.*, Observation of reactor antineutrino disappearance using delayed neutron capture on hydrogen at RENO, *J. High Energ. Phys.*, 2020 2020(4) 29, http://dx.doi.org/10.1007/JHEP04(2020)029

[41] M. A. Acero *et al.*, An Improved Measurement of Neutrino Oscillation Parameters by the NOvA Experiment, *arXiv:2108.08219 [hep-ex]*, 2021 ArXiv: 2108.08219

[42] L. Wolfenstein, Neutrino oscillations in matter, *Phys. Rev. D*, 1978 17(9) 2369, http://dx.doi.org/10.1103/PhysRevD.17.2369, publisher: American Physical Society

[43] S. P. Mikheyev A. Y. Smirnov, Resonant amplification of neutrino oscillations in matter and solar-neutrino spectroscopy, *Il Nuovo Cimento C*, 1986 9(1) 17, http://dx.doi.org/10.1007/BF02508049

[44] J. N. Bahcall C. Peña-Garay, A road map to solar neutrino fluxes, neutrino oscillation parameters, and tests for new physics, *J. High Energy Phys.*, 2003 2003(11) 004, http://dx.doi.org/10.1088/1126-6708/2003/11/004, publisher: Springer Science and Business Media LLC

[45] G. Fogli E. Lisi, Evidence for the MSW effect, *New J. Phys.*, 2004 6 139, http://dx.doi.org/10.1088/1367-2630/6/1/139

[46] The Borexino Collaboration, Comprehensive measurement of pp-chain solar neutrinos, *Nature*, 2018 562(7728) 505, http://dx.doi.org/10.1038/s41586-018-0624-y

[47] Y. Nakajima, Recent results and future prospects from Super-Kamiokande, 2020, http://dx.doi.org/10.5281/zenodo.3959640

[48] S. Turck-Chièze, The Standard Solar Model and beyond, *J. Phys.: Conf. Ser.*, 2016 665 012078, http://dx.doi.org/10.1088/1742-6596/665/1/012078, publisher: IOP Publishing

[49] D. A. Rothery *et al.*, Our worlds. the magnetism and thrill of planetary exploration, *Geological Magazine*, 2000 137(4) 463, http://dx.doi.org/10.1017/S0016756800234419, publisher: Cambridge University Press

[50] E. Caffau *et al.*, The solar photospheric nitrogen abundance - Analysis of atomic transitions with 3D and 1D model atmospheres, *A&A*, 2009 498(3) 877, http://dx.doi.org/10.1051/0004-6361/200810859, number: 3 Publisher: EDP Sciences

[51] S. Turck-Chièze *et al.*, Surprising Sun: A New Step Towards a Complete Picture?, *Phys. Rev. Lett.*, 2004 93(21) 211102, http://dx.doi.org/10.1103/PhysRevLett.93.211102, publisher: American Physical Society

[52] N. Grevesse A. Sauval, Standard Solar Composition, *Space Science Reviews*, 1998 85(1) 161, http://dx.doi.org/10.1023/A:1005161325181

[53] J. N. Bahcall *et al.*, Helioseismological Implications of Recent Solar Abundance Determinations, *ApJ*, 2005 618(2) 1049, http://dx.doi.org/10.1086/426070, publisher: IOP Publishing

[54] J. Christensen-Dalsgaard *et al.*, On the opacity change required to compensate for the revised solar composition, *A&A*, 2009 494(1) 205, http://dx.doi.org/10.1051/0004-6361:200810170, number: 1 Publisher: EDP Sciences

[55] F. L. Villante, Constraints on the opacity profile of the sun from helioseismic observables and solar neutrino flux measurements, *ApJ*, 2010 724(1) 98, http://dx.doi.org/10.1088/0004-637X/724/1/98, publisher: American Astronomical Society

[56] H. A. Bethe, Energy Production in Stars, *Phys. Rev.*, 1939 55(1) 103, http://dx.doi.org/10.1103/PhysRev.55.103, publisher: American Physical Society

[57] C. F. Weizsäcker, Über Elementumwandlungen in Innern der Sterne, *Physikalische Zeitschrift*, 1937 38 176

[58] W. C. Haxton A. M. Serenelli, CN Cycle Solar Neutrinos and the Sun's Primordial Core Metallicity, *ApJ*, 2008 687(1) 678, http://dx.doi.org/10.1086/591787, publisher: IOP Publishing

[59] A. Serenelli, C. Peña-Garay W. C. Haxton, Using the standard solar model to constrain solar composition and nuclear reaction S factors, *Phys. Rev. D*, 2013 87(4) 043001, http://dx.doi.org/10.1103/PhysRevD.87.043001

[60] BOREXINO Collaboration *et al.*, Sensitivity to neutrinos from the solar CNO cycle in Borexino, *Eur. Phys. J. C*, 2020 80(11) 1091, http://dx.doi.org/10.1140/epjc/s10052-020-08534-2

[61] R. d. E. Atkinson F. G. Houtermans, Zur Frage der Aufbaumöglichkeit der Elemente in Sternen, *Z. Physik*, 1929 54(9) 656, http://dx.doi.org/10.1007/BF01341595

[62] H. A. Bethe C. L. Critchfield, The Formation of Deuterons by Proton Combination, *Phys. Rev.*, 1938 54(4) 248, http://dx.doi.org/10.1103/PhysRev.54.248, publisher: American Physical Society

[63] W. A. Fowler, Experimental and theoretical nuclear astrophysics: the quest for the origin of the elements, *Rev. Mod. Phys.*, 1984 56(2) 149, http://dx.doi.org/10.1103/RevModPhys.56.149, publisher: American Physical Society

[64] R. Davis, Nobel Lecture: A half-century with solar neutrinos, *Rev. Mod. Phys.*, 2003 75(3) 985, http://dx.doi.org/10.1103/RevModPhys.75.985, publisher: American Physical Society

[65] K. S. Hirata *et al.*, Observation of 8b solar neutrinos in the Kamiokande-II detector, *Phys. Rev. Lett.*, 1989 63(1) 16, http://dx.doi.org/10.1103/PhysRevLett.63.16, publisher: American Physical Society

[66] M. Salaris S. Cassisi, Evolution of Stars and Stellar Populations, 2005

[67] R. Kippenhahn, A. Weigert A. Weiss, Stellar Structure and Evolution, R. Kippenhahn, A. Weigert A. Weiss, eds., Stellar Structure and Evolution, Astronomy and Astrophysics Library, Springer, Berlin, Heidelberg, 2012 73–81, http://dx.doi.org/10.1007/978-3-642-30304-3_8

[68] M. Agostini *et al.*, Experimental evidence of neutrinos produced in the CNO fusion cycle in the Sun, *Nature*, 2020 587(7835) 577, http://dx.doi.org/10.1038/s41586-020-2934-0, number: 7835 Publisher: Nature Publishing Group

[69] S. C. Schuler, J. R. King L.-S. The, Stellar nucleosynthesis in the hyades open cluster, *ApJ*, 2009 701(1) 837, http://dx.doi.org/10.1088/0004-637X/701/1/837, publisher: American Astronomical Society

[70] A. Mesa-Delgado *et al.*, The trace of the cno cycle in the ring nebular ngc 6888, *ApJ*, 2014 785(2) 100, http://dx.doi.org/10.1088/0004-637X/785/2/100, publisher: American Astronomical Society

[71] X. Qian J.-C. Peng, Physics with reactor neutrinos, *Rep. Prog. Phys.*, 2019 82(3) 036201, http://dx.doi.org/10.1088/1361-6633/aae881, publisher: IOP Publishing

[72] M. Sajjad Athar et al., The IUPAP Neutrino Panel Report, 2021

[73] Daya Bay Collaboration *et al.*, Improved measurement of the reactor antineutrino flux at Daya Bay, *Phys. Rev. D*, 2019 100(5) 052004, http://dx.doi.org/10.1103/PhysRevD.100.052004, publisher: American Physical Society

[74] The KamLAND Collaboration *et al.*, Constraints on theta13 from a three-flavor oscillation analysis of reactor antineutrinos at KamLAND, *Phys. Rev. D*, 2011 83(5) 052002, http://dx.doi.org/10.1103/PhysRevD.83.052002, publisher: American Physical Society

[75] T. A. Mueller *et al.*, Improved predictions of reactor antineutrino spectra, *Phys. Rev. C*, 2011 83(5) 054615, http://dx.doi.org/10.1103/PhysRevC.83.054615, publisher: American Physical Society

[76] RENO Collaboration *et al.*, Measurement of reactor antineutrino flux and spectrum at RENO, *Phys. Rev. D*, 2021 104(11) L111301, http://dx.doi.org/10.1103/PhysRevD.104.L111301, publisher: American Physical Society

[77] Daya Bay Collaboration *et al.*, Evolution of the Reactor Antineutrino Flux and Spectrum at Daya Bay, *Phys. Rev. Lett.*, 2017 118(25) 251801, http://dx.doi.org/10.1103/PhysRevLett.118.251801, publisher: American Physical Society

[78] RENO Collaboration *et al.*, Fuel-Composition Dependent Reactor Antineutrino Yield at RENO, *Phys. Rev. Lett.*, 2019 122(23) 232501, http://dx.doi.org/10.1103/PhysRevLett.122.232501, publisher: American Physical Society

[79] V. Kopeikin, M. Skorokhvatov O. Titov, Reevaluating reactor antineutrino spectra with new measurements of the ratio between 235-u and 239-pu beta spectra, *Phys. Rev. D*, 2021 104(7) L071301, http://dx.doi.org/10.1103/PhysRevD.104.L071301, publisher: American Physical Society

[80] STEREO Collaboration *et al.*, Improved sterile neutrino constraints from the STEREO experiment with 179 days of reactor-on data, *Phys. Rev. D*, 2020 102(5) 052002, http://dx.doi.org/10.1103/PhysRevD.102.052002, publisher: American Physical Society

[81] S.-H. Seo, New results from RENO and the 5 MeV excess, *AIP Conference Proceedings*, 2015 1666(1) 080002, http://dx.doi.org/10.1063/1.4915563, publisher: American Institute of Physics

[82] Y. Abe *et al.*, Erratum to: Improved measurements of the neutrino mixing angle theta13 with the Double Chooz detector, *J. High Energ. Phys.*, 2015 2015(2) 74, http://dx.doi.org/10.1007/JHEP02(2015)074

[83] Anatael Cabrera, Double Chooz Gd-III: First results, 2014

[84] Daya Bay Collaboration *et al.*, Measurement of the Reactor Antineutrino Flux and Spectrum at Daya Bay, *Phys. Rev. Lett.*, 2016 116(6) 061801, http://dx.doi.org/10.1103/PhysRevLett.116.061801, publisher: American Physical Society

[85] Daya Bay Collaboration *et al.*, Extraction of the 235u and 239pu Antineutrino Spectra at Daya Bay, *Phys. Rev. Lett.*, 2019 123(11) 111801, http://dx.doi.org/10.1103/PhysRevLett.123.111801, publisher: American Physical Society

[86] V. A. Balkanov *et al.*, Registration of atmospheric neutrinos with the BAIKAL Neutrino Telescope NT-96, *Astroparticle Physics*, 1999 12(1) 75, http://dx.doi.org/10.1016/S0927-6505(99)00078-X

[87] M. G. Aartsen *et al.*, Time-Integrated Neutrino Source Searches with 10 Years of IceCube Data, *Phys. Rev. Lett.*, 2020 124(5) 051103, http://dx.doi.org/10.1103/PhysRevLett.124.051103, publisher: American Physical Society

[88] A. Sinopoulou *et al.*, Atmospheric neutrinos with the first detection units of KM3NeT/ARCA, *J. Inst.*, 2021 16(11) C11015, http://dx.doi.org/10.1088/1748-0221/16/11/C11015, publisher: IOP Publishing

[89] A. Albert *et al.*, Measurement of the atmospheric nu_e and nu_mu energy spectra with the ANTARES neutrino telescope, *Physics Letters B*, 2021 816 136228, http://dx.doi.org/10.1016/j.physletb.2021.136228

[90] Super-Kamiokande Collaboration *et al.*, Measurement of the tau neutrino cross section in atmospheric neutrino oscillations with Super-Kamiokande, *Phys. Rev. D*, 2018 98(5) 052006, http://dx.doi.org/10.1103/PhysRevD.98.052006, publisher: American Physical Society

[91] IceCube Collaboration 1 *et al.*, Measurement of atmospheric tau neutrino appearance with IceCube DeepCore, *Phys. Rev. D*, 2019 99(3) 032007, http://dx.doi.org/10.1103/PhysRevD.99.032007, publisher: American Physical Society

[92] F. An *et al.*, Neutrino physics with JUNO, *J. Phys. G: Nucl. Part. Phys.*, 2016 43(3) 030401, http://dx.doi.org/10.1088/0954-3899/43/3/030401, publisher: IOP Publishing

[93] M. G. Aartsen *et al.*, PINGU: a vision for neutrino and particle physics at the South Pole, *J. Phys. G: Nucl. Part. Phys.*, 2017 44(5) 054006, http://dx.doi.org/10.1088/1361-6471/44/5/054006, publisher: IOP Publishing

[94] IceCube-Gen2 Collaboration *et al.*, Combined sensitivity to the neutrino mass ordering with JUNO, the IceCube Upgrade, and PINGU, *Phys. Rev. D*, 2020 101(3) 032006, http://dx.doi.org/10.1103/PhysRevD.101.032006, publisher: American Physical Society

[95] H.-K. Proto-Collaboration *et al.*, Hyper-Kamiokande Design Report, *arXiv:1805.04163 [astro-ph, physics:hep-ex, physics:physics]*, 2018 ArXiv: 1805.04163

[96] S. Aiello *et al.*, Determining the neutrino mass ordering and oscillation parameters with KM3NeT/ORCA, *Eur. Phys. J. C*, 2022 82(1) 26, http://dx.doi.org/10.1140/epjc/s10052-021-09893-0

[97] H. A. Bethe J. R. Wilson, Revival of a stalled supernova shock by neutrino heating, *The Astrophysical Journal*, 1985 295 14, http://dx.doi.org/10.1086/163343, aDS Bibcode: 1985ApJ...295...14B

[98] A. Burrows, Supernova explosions in the Universe, *Nature*, 2000 403(6771) 727, http://dx.doi.org/10.1038/35001501, number: 6771 Publisher: Nature Publishing Group

[99] H.-T. Janka, Explosion Mechanisms of Core-Collapse Supernovae, *Annual Review of Nuclear and Particle Science*, 2012 62 407, http://dx.doi.org/10.1146/annurev-nucl-102711-094901, aDS Bibcode: 2012ARNPS..62..407J

[100] K. Hirata *et al.*, Observation of a neutrino burst from the supernova SN1987A, *Phys. Rev. Lett.*, 1987 58(14) 1490, http://dx.doi.org/10.1103/PhysRevLett.58.1490, publisher: American Physical Society

[101] JETP Letters: issues online

[102] R. M. Bionta *et al.*, Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud, *Phys. Rev. Lett.*, 1987 58(14) 1494, http://dx.doi.org/10.1103/PhysRevLett.58.1494, publisher: American Physical Society

[103] S. A. Kharusi *et al.*, SNEWS 2.0: a next-generation supernova early warning system for multi-messenger astronomy, *New J. Phys.*, 2021 23(3) 031201, http://dx.doi.org/10.1088/1367-2630/abde33, publisher: IOP Publishing

[104] J.-S. Lu *et al.*, Constraining absolute neutrino masses via detection of galactic supernova neutrinos at JUNO, *J. Cosmol. Astropart. Phys.*, 2015 2015(05) 044, http://dx.doi.org/10.1088/1475-7516/2015/05/044, publisher: IOP Publishing

[105] K. MØller *et al.*, Measuring the supernova unknowns at the next-generation neutrino telescopes through the diffuse neutrino background, *J. Cosmol. Astropart. Phys.*, 2018 2018(05) 066, http://dx.doi.org/10.1088/1475-7516/2018/05/066, publisher: IOP Publishing

[106] A. Abusleme *et al.*, JUNO sensitivity to low energy atmospheric neutrino spectra, *Eur. Phys. J. C*, 2021 81(10) 887, http://dx.doi.org/10.1140/epjc/s10052-021-09565-z

[107] G. Fiorentini, M. Lissia F. Mantovani, Geo-neutrinos and earth's interior, *Physics Reports*, 2007 453(5) 117, http://dx.doi.org/10.1016/j.physrep.2007.09.001

[108] A. Rocholl K. P. Jochum, Th, U and other trace elements in carbonaceous chondrites: Implications for the terrestrial and solar-systemTh/U ratios, *Earth*

and Planetary Science Letters, 1993 117(1) 265, http://dx.doi.org/10.1016/0012-821X(93)90132-S

[109] H. Watanabe, Geo-neutrinos, 2020, http://dx.doi.org/10.5281/zenodo.3959690

[110] Borexino Collaboration *et al.*, Comprehensive geoneutrino analysis with Borexino, *Phys. Rev. D*, 2020 101(1) 012009, http://dx.doi.org/10.1103/PhysRevD.101.012009, publisher: American Physical Society

[111] R. Han *et al.*, Potential of geo-neutrino measurements at JUNO, *Chinese Phys. C*, 2016 40(3) 033003, http://dx.doi.org/10.1088/1674-1137/40/3/033003, publisher: IOP Publishing

[112] M. C. Chen, Geo-neutrinos in SNO+, *Earth Moon Planet*, 2006 99(1) 221, http://dx.doi.org/10.1007/s11038-006-9116-4

[113] L. Wan *et al.*, Geoneutrinos at Jinping: Flux prediction and oscillation analysis, *Phys. Rev. D*, 2017 95(5) 053001, http://dx.doi.org/10.1103/PhysRevD.95.053001, publisher: American Physical Society

[114] I. R. Barabanov *et al.*, Large-volume detector at the Baksan Neutrino Observatory for studies of natural neutrino fluxes for purposes of geo- and astrophysics, *Phys. Atom. Nuclei*, 2017 80(3) 446, http://dx.doi.org/10.1134/S1063778817030036

[115] M. Schwartz, Feasibility of Using High-Energy Neutrinos to Study the Weak Interactions, *Phys. Rev. Lett.*, 1960 4(6) 306, http://dx.doi.org/10.1103/PhysRevLett.4.306, publisher: American Physical Society

[116] U. Dore, P. Loverre L. Ludovici, History of accelerator neutrino beams, *EPJ H*, 2019 44(4) 271, http://dx.doi.org/10.1140/epjh/e2019-90032-x

[117] S. Cao *et al.*, Physics potential of the combined sensitivity of T2K-II, nova extension, and juno, *Phys. Rev. D*, 2021 103(11) 112010, http://dx.doi.org/10.1103/PhysRevD.103.112010, publisher: American Physical Society

[118] J. Cao *et al.*, On the complementarity of Hyper-K and LBNF, *arXiv:1501.03918 [hep-ex, physics:physics]*, 2015 ArXiv: 1501.03918

[119] M. Kuze, Neutrino Experiments at J-PARC, Proceedings of the 3rd J-PARC Symposium (J-PARC2019), *JPS Conference Proceedings*, vol. 33, Journal of the Physical Society of Japan, 2021 http://dx.doi.org/10.7566/JPSCP.33.011139

[120] D. Naples *et al.*, High energy neutrino scattering results from NuTeV, *Nuclear Physics B - Proceedings Supplements*, 2003 118 164, http://dx.doi.org/10.1016/S0920-5632(03)01314-8

[121] M. Muether, NOvA: Current Status and Future Reach, *Nuclear Physics B - Proceedings Supplements*, 2013 237-238 135, http://dx.doi.org/10.1016/j.nuclphysbps.2013.04.075

[122] MiniBooNE Collaboration *et al.*, First Measurement of Monoenergetic Muon Neutrino Charged Current Interactions, *Phys. Rev. Lett.*, 2018 120(14) 141802, http://dx.doi.org/10.1103/PhysRevLett.120.141802, publisher: American Physical Society

[123] A. De Rújula, M. B. Gavela P. Hernández, Neutrino oscillation physics with a neutrino factory, *Nuclear Physics B*, 1999 547(1) 21, http://dx.doi.org/10.1016/S0550-3213(99)00070-X

[124] J. B. Birks, CHAPTER 5 - THE DETECTION OF SCINTILLATIONS, J. B. Birks, ed., The Theory and Practice of Scintillation Counting, International Series of Monographs in Electronics and Instrumentation, Pergamon, 1964 96–184, http://dx.doi.org/10.1016/B978-0-08-010472-0.50010-0

[125] J. B. Birks, Scintillations from Organic Crystals: Specific Fluorescence and Relative Response to Different Radiations, *Proc. Phys. Soc. A*, 1951 64(10) 874, http://dx.doi.org/10.1088/0370-1298/64/10/303, publisher: IOP Publishing

[126] Borexino Collaboration *et al.*, Correlated and integrated directionality for sub-MeV solar neutrinos in Borexino, *Phys. Rev. D*, 2022 105(5) 052002, http://dx.doi.org/10.1103/PhysRevD.105.052002, publisher: American Physical Society

[127] Borexino Collaboration *et al.*, First Directional Measurement of Sub-MeV Solar Neutrinos with Borexino, *Phys. Rev. Lett.*, 2022 128(9) 091803, http://dx.doi.org/10.1103/PhysRevLett.128.091803, publisher: American Physical Society

[128] M. Askins *et al.*, Theia: an advanced optical neutrino detector, *Eur. Phys. J. C*, 2020 80(5) 416, http://dx.doi.org/10.1140/epjc/s10052-020-7977-8

[129] G. Ranucci, A. Goretti P. Lombardi, Pulse-shape discrimination of liquid scintillators, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1998 412(2) 374, http://dx.doi.org/10.1016/S0168-9002(98)00456-2

[130] H. Rebber *et al.*, Particle identification at MeV energies in JUNO, *J. Inst.*, 2021 16(01) P01016, http://dx.doi.org/10.1088/1748-0221/16/01/P01016

[131] Atlas of Neutron Capture Cross Sections, 2019, publisher: IAEA

[132] J. Collaboration *et al.*, TAO Conceptual Design Report: A Precision Measurement of the Reactor Antineutrino Spectrum with Sub-percent Energy Resolution, *arXiv:2005.08745 [hep-ex, physics:nucl-ex, physics:physics]*, 2020 ArXiv: 2005.08745

[133] RENO Collaboration *et al.*, Measurement of Reactor Antineutrino Oscillation Amplitude and Frequency at RENO, *Phys. Rev. Lett.*, 2018 121(20) 201801, http://dx.doi.org/10.1103/PhysRevLett.121.201801, publisher: American Physical Society

[134] H. J. Ache, C. I. of Canada A. C. Society, eds., Positronium and muonium chemistry, no. 175 in Advances in chemistry series, American Chemical Society, Washington, 1979

[135] D. Franco, G. Consolati D. Trezzi, Positronium signature in organic liquid scintillators for neutrino experiments, *Phys. Rev. C*, 2011 83(1) 015504, http://dx.doi.org/10.1103/PhysRevC.83.015504, publisher: American Physical Society

[136] Y. Abe *et al.*, Ortho-positronium observation in the Double Chooz experiment, *J. High Energ. Phys.*, 2014 2014(10) 32, http://dx.doi.org/10.1007/JHEP10(2014)032

[137] Borexino Collaboration *et al.*, Simultaneous precision spectroscopy of pp, 7Be, and pep solar neutrinos with {Borexino} {Phase}-{II}, *Phys. Rev. D*, 2019 100(8) 082004, http://dx.doi.org/10.1103/PhysRevD.100.082004, publisher: American Physical Society

[138] O. Smirnov, Experimental aspects of geoneutrino detection: Status and perspectives, *Progress in Particle and Nuclear Physics*, 2019 109 103712, http://dx.doi.org/10.1016/j.ppnp.2019.103712

[139] J. Bernabéu, S. Palomares-Ruiz S. T. Petcov, Atmospheric neutrino oscillations, theta13 and neutrino mass hierarchy, *Nuclear Physics B*, 2003 669(1) 255, http://dx.doi.org/10.1016/j.nuclphysb.2003.07.025

[140] A. S. Dighe A. Y. Smirnov, Identifying the neutrino mass spectrum from a supernova neutrino burst, *Phys. Rev. D*, 2000 62(3) 033007, http://dx.doi.org/10.1103/PhysRevD.62.033007, publisher: American Physical Society

[141] M. Fukugita T. Yanagida, Barygenesis without grand unification, *Physics Letters B*, 1986 174(1) 45, http://dx.doi.org/10.1016/0370-2693(86)91126-3

[142] W. Buchmüller, P. Di Bari M. Plümacher, Leptogenesis for pedestrians, *Annals of Physics*, 2005 315(2) 305, http://dx.doi.org/10.1016/j.aop.2004.02.003

[143] S. Davidson, E. Nardi Y. Nir, Leptogenesis, *Physics Reports*, 2008 466(4) 105, http://dx.doi.org/10.1016/j.physrep.2008.06.002

[144] C. Palomares, DUNE Physics Program and Status, *J. Phys.: Conf. Ser.*, 2021 2156(1) 012107, http://dx.doi.org/10.1088/1742-6596/2156/1/012107, publisher: IOP Publishing

[145] P. Ramond, The family group in grand unified theories, SEESAW 25, WORLD SCIENTIFIC, 2005 265–280, http://dx.doi.org/10.1142/9789812702210_0020

[146] M. Gell-Mann, P. Ramond R. Slansky, Complex spinors and unified theories, Murray Gell-Mann, *World Scientific Series in 20th Century Physics*, vol. Volume 40, WORLD SCIENTIFIC, 2010 266–272, http://dx.doi.org/10.1142/9789812836854_0018

[147] A. A. Esfahani *et al.*, Determining the neutrino mass with cyclotron radiation emission spectroscopy—Project 8, *J. Phys. G: Nucl. Part. Phys.*, 2017 44(5) 054004, http://dx.doi.org/10.1088/1361-6471/aa5b4f, publisher: IOP Publishing

[148] B. Monreal J. A. Formaggio, Relativistic cyclotron radiation detection of tritium decay electrons as a new technique for measuring the neutrino mass, *Phys. Rev. D*, 2009 80(5) 051301, http://dx.doi.org/10.1103/PhysRevD.80.051301, publisher: American Physical Society

[149] A. Nucciotti *et al.*, Status of the HOLMES Experiment to Directly Measure the Neutrino Mass, *J Low Temp Phys*, 2018 193(5) 1137, http://dx.doi.org/10.1007/s10909-018-2025-x

[150] A. De Rújula M. Lusignoli, Calorimetric measurements of 163holmium decay as tools to determine the electron neutrino mass, *Physics Letters B*, 1982 118(4) 429, http://dx.doi.org/10.1016/0370-2693(82)90218-0

[151] E. Di Valentino, S. Gariazzo O. Mena, Most constraining cosmological neutrino mass bounds, *Phys. Rev. D*, 2021 104(8) 083504, http://dx.doi.org/10.1103/PhysRevD.104.083504, publisher: American Physical Society

[152] S. Dell'Oro *et al.*, Neutrinoless Double Beta Decay: 2015 Review, *Advances in High Energy Physics*, 2016 2016 e2162659, http://dx.doi.org/10.1155/2016/2162659, publisher: Hindawi

[153] J. Engel J. Menéndez, Status and future of nuclear matrix elements for neutrinoless double-beta decay: a review, *Rep. Prog. Phys.*, 2017 80(4) 046301, http://dx.doi.org/10.1088/1361-6633/aa5bc5, publisher: IOP Publishing

[154] KamLAND-Zen Collaboration *et al.*, Search for Majorana Neutrinos Near the Inverted Mass Hierarchy Region with KamLAND-Zen, *Phys. Rev. Lett.*, 2016 117(8) 082503, http://dx.doi.org/10.1103/PhysRevLett.117.082503, publisher: American Physical Society

[155] GERDA Collaboration *et al.*, Final resutls of gerda on the search for neutrinoless double-beta decay, *Phys. Rev. Lett.*, 2020 125(25) 252502, http://dx.doi.org/10.1103/PhysRevLett.125.252502, publisher: American Physical Society

[156] A. Abusleme et al., JUNO physics and detector, *Progress in Particle and Nuclear Physics*, 2022 123 103927, http://dx.doi.org/10.1016/j.ppnp.2021.103927

[157] LSND Collaboration *et al.*, Evidence for neutrino oscillations from the observation of anti˙nu˙e appearance in a anti˙nu˙mu beam, *Phys. Rev. D*, 2001 64(11) 112007, http://dx.doi.org/10.1103/PhysRevD.64.112007, publisher: American Physical Society

[158] MiniBooNE Collaboration *et al.*, Significant Excess of Electronlike Events in the MiniBooNE Short-Baseline Neutrino Experiment, *Phys. Rev. Lett.*, 2018 121(22) 221801, http://dx.doi.org/10.1103/PhysRevLett.121.221801, publisher: American Physical Society

[159] C. Giunti M. Laveder, Statistical significance of the gallium anomaly, *Phys. Rev. C*, 2011 83(6) 065504, http://dx.doi.org/10.1103/PhysRevC.83.065504, publisher: American Physical Society

[160] V. V. Barinov *et al.*, Results from the Baksan Experiment on Sterile Transitions (BEST), *arXiv:2109.11482 [hep-ex, physics:nucl-ex]*, 2021 ArXiv: 2109.11482

[161] NOvA Collaboration *et al.*, Constraints on oscillation parameters from nu˙e appearance and nu˙mu disappearance in nova, *Phys. Rev. Lett.*, 2017 118(23) 231801, http://dx.doi.org/10.1103/PhysRevLett.118.231801, publisher: American Physical Society

[162] M. Dentler *et al.*, Updated global analysis of neutrino oscillations in the presence of eV-scale sterile neutrinos, *J. High Energ. Phys.*, 2018 2018(8) 10, http://dx.doi.org/10.1007/JHEP08(2018)010

[163] R. Acciarri *et al.*, Design and construction of the MicroBooNE detector, *J. Inst.*, 2017 12(02) P02017, http://dx.doi.org/10.1088/1748-0221/12/02/P02017, publisher: IOP Publishing

[164] R. Acciarri *et al.*, A Proposal for a Three Detector Short-Baseline Neutrino Oscillation Program in the Fermilab Booster Neutrino Beam, *arXiv:1503.01520 [hep-ex, physics:physics]*, 2015 ArXiv: 1503.01520

[165] A. P. Serebrov *et al.*, First Observation of the Oscillation Effect in the Neutrino-4 Experiment on the Search for the Sterile Neutrino, *Jetp Lett.*, 2019 109(4) 213, http://dx.doi.org/10.1134/S0021364019040040

[166] Daya Bay Collaboration *et al.*, Improved Constraints on Sterile Neutrino Mixing from Disappearance Searches in the MINOS, minos+, Daya Bay, and Bugey-3 Experiments, *Phys. Rev. Lett.*, 2020 125(7) 071801, http://dx.doi.org/10.1103/PhysRevLett.125.071801, publisher: American Physical Society

[167] KamLAND Collaboration *et al.*, 7be solar neutrino measurement with KamLAND, *Phys. Rev. C*, 2015 92(5) 055808, http://dx.doi.org/10.1103/PhysRevC.92.055808, publisher: American Physical Society

[168] D. Svirida, DANSS experiment: current status and future plans, *J. Phys.: Conf. Ser.*, 2020 1690(1) 012179, http://dx.doi.org/10.1088/1742-6596/1690/1/012179, publisher: IOP Publishing

[169] M. Andriamirado *et al.*, PROSPECT-II Physics Opportunities, *arXiv:2107.03934 [hep-ex, physics:nucl-ex, physics:physics]*, 2021 ArXiv: 2107.03934

[170] S. Böser *et al.*, Status of light sterile neutrino searches, *Progress in Particle and Nuclear Physics*, 2020 111 103736, http://dx.doi.org/10.1016/j.ppnp.2019.103736

[171] Y. Farzan M. Tórtola, Neutrino Oscillations and Non-standard Interactions, *Frontiers in Physics*, 2018 6

[172] I. Esteban *et al.*, Updated constraints on non-standard interactions from global analysis of oscillation data, *J. High Energ. Phys.*, 2018 2018(8) 180, http://dx.doi.org/10.1007/JHEP08(2018)180

[173] Y. Farzan I. M. Shoemaker, Lepton flavor violating non-standard interactions via light mediators, *J. High Energ. Phys.*, 2016 2016(7) 33, http://dx.doi.org/10.1007/JHEP07(2016)033

[174] C. Giunti A. Studenikin, Neutrino electromagnetic properties, *Phys. Atom. Nuclei*, 2009 72(12) 2089, http://dx.doi.org/10.1134/S1063778809120126

[175] G. G. Raffelt, Limits on neutrino electromagnetic properties — an update, *Physics Reports*, 1999 320(1) 319, http://dx.doi.org/10.1016/S0370-1573(99)00074-5

[176] C. Arpesella *et al.*, First real time detection of 7Be solar neutrinos by Borexino, *Physics Letters B*, 2008 658(4) 101, http://dx.doi.org/10.1016/j.physletb.2007.09.054

[177] Borexino Collaboration *et al.*, Measurement of the solar 8b neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector, *Phys. Rev. D*, 2010 82(3) 033006, http://dx.doi.org/10.1103/PhysRevD.82.033006, publisher: American Physical Society

[178] Super-Kamiokande Collaboration *et al.*, Solar ^{8}B and hep Neutrino Measurements from 1258 Days of Super-Kamiokande Data, *Phys. Rev. Lett.*, 2001 86(25) 5651, http://dx.doi.org/10.1103/PhysRevLett.86.5651, publisher: American Physical Society

[179] Borexino Collaboration *et al.*, Precision Measurement of the 7Be Solar Neutrino Interaction Rate in Borexino, *Phys. Rev. Lett.*, 2011 107(14) 141302, http://dx.doi.org/10.1103/PhysRevLett.107.141302, publisher: American Physical Society

[180] Borexino Collaboration *et al.*, First Evidence of pep Solar Neutrinos by Direct Detection in Borexino, *Phys. Rev. Lett.*, 2012 108(5) 051302, http://dx.doi.org/10.1103/PhysRevLett.108.051302, publisher: American Physical Society

[181] Borexino Collaboration *et al.*, Final results of Borexino Phase-I on low-energy solar neutrino spectroscopy, *Phys. Rev. D*, 2014 89(11) 112007, http://dx.doi.org/10.1103/PhysRevD.89.112007, publisher: American Physical Society

[182] G. Bellini *et al.*, Observation of geo-neutrinos, *Physics Letters B*, 2010 687(4) 299, http://dx.doi.org/10.1016/j.physletb.2010.03.051

[183] G. Bellini *et al.*, Neutrinos from the primary proton–proton fusion process in the Sun, *Nature*, 2014 512(7515) 383, http://dx.doi.org/10.1038/nature13702, number: 7515 Publisher: Nature Publishing Group

[184] D. Bravo-Berguño *et al.*, The Borexino Thermal Monitoring & Management System and simulations of the fluid-dynamics of the Borexino detector under asymmetrical, changing boundary conditions, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2018 885 38, http://dx.doi.org/10.1016/j.nima.2017.12.047

[185] The Borexino collaboration *et al.*, Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data, *Phys. Rev. D*, 2017 96(9) 091103, http://dx.doi.org/10.1103/PhysRevD.96.091103, publisher: American Physical Society

[186] M. Agostini *et al.*, Modulations of the cosmic muon signal in ten years of Borexino data, *J. Cosmol. Astropart. Phys.*, 2019 2019(02) 046, http://dx.doi.org/10.1088/1475-7516/2019/02/046, publisher: IOP Publishing

[187] S. K. Agarwalla *et al.*, Constraints on flavor-diagonal non-standard neutrino interactions from Borexino Phase-II, *J. High Energ. Phys.*, 2020 2020(2) 38, http://dx.doi.org/10.1007/JHEP02(2020)038

[188] S. Appel *et al.*, Search for low-energy signals from fast radio bursts with the Borexino detector, *Eur. Phys. J. C*, 2022 82(3) 278, http://dx.doi.org/10.1140/epjc/s10052-022-10197-0

[189] M. Redchuk, Looking inside the sun with the Borexino experiment: detection of solar neutrinos from the proton-proton chain and the CNO cycle, Ph.D. thesis, RWTH Aachen University, 2020, http://dx.doi.org/10.18154/RWTH-2020-12244, number: RWTH-2020-12244

[190] G. Bellini *et al.*, Muon and cosmogenic neutron detection in Borexino, *J. Inst.*, 2011 6(05) P05005, http://dx.doi.org/10.1088/1748-0221/6/05/P05005, publisher: IOP Publishing

[191] M. Ambrosio *et al.*, Measurement of the residual energy of muons in the Gran Sasso underground laboratories, *Astroparticle Physics*, 2003 19(3) 313, http://dx.doi.org/10.1016/S0927-6505(02)00217-7

[192] H. Back *et al.*, Borexino calibrations: hardware, methods, and results, *J. Inst.*, 2012 7(10) P10018, http://dx.doi.org/10.1088/1748-0221/7/10/P10018, publisher: IOP Publishing

[193] A. Abusleme *et al.*, Optimization of the JUNO liquid scintillator composition using a Daya Bay antineutrino detector, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2021 988 164823, http://dx.doi.org/10.1016/j.nima.2020.164823

[194] A. Abusleme *et al.*, Calibration strategy of the JUNO experiment, *J. High Energ. Phys.*, 2021 2021(3) 4, http://dx.doi.org/10.1007/JHEP03(2021)004

[195] D. A. Dwyer T. J. Langford, Spectral Structure of Electron Antineutrinos from Nuclear Reactors, *Phys. Rev. Lett.*, 2015 114(1) 012502, http://dx.doi.org/10.1103/PhysRevLett.114.012502, publisher: American Physical Society

[196] A. Abusleme *et al.*, Feasibility and physics potential of detecting 8B solar neutrinos at JUNO \ast, *Chinese Phys. C*, 2021 45(2) 023004, http://dx.doi.org/10.1088/1674-1137/abd92a, publisher: IOP Publishing

[197] I. Esteban *et al.*, Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of theta23, deltaCP, and the mass ordering, *J. High Energ. Phys.*, 2019 2019(1) 106, http://dx.doi.org/10.1007/JHEP01(2019)106

[198] KamLAND Collaboration *et al.*, Measurement of the ^{8}B solar neutrino flux with the KamLAND liquid scintillator detector, *Phys. Rev. C*, 2011 84(3) 035804, http://dx.doi.org/10.1103/PhysRevC.84.035804, publisher: American Physical Society

[199] M. Reguzzoni *et al.*, GIGJ: A Crustal Gravity Model of the Guangdong Province for Predicting the Geoneutrino Signal at the JUNO Experiment, *Journal of Geophysical Research: Solid Earth*, 2019 124(4) 4231, http://dx.doi.org/10.1029/2018JB016681, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018JB016681

[200] R. Gao *et al.*, JULOC: A local 3-D high-resolution crustal model in South China for forecasting geoneutrino measurements at JUNO, *Physics of the Earth and Planetary Interiors*, 2020 299 106409, http://dx.doi.org/10.1016/j.pepi.2019.106409

[201] V. Strati *et al.*, Expected geoneutrino signal at JUNO, *arXiv:1412.3324 [physics]*, 2015 ArXiv: 1412.3324

[202] J. Liu *et al.*, Automated calibration system for a high-precision measurement of neutrino mixing angle theta13 with the Daya Bay antineutrino detectors, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2014 750 19, http://dx.doi.org/10.1016/j.nima.2014.02.049

[203] A. Abusleme *et al.*, Radioactivity control strategy for the JUNO detector, *J. High Energ. Phys.*, 2021 2021(11) 102, http://dx.doi.org/10.1007/JHEP11(2021)102

[204] G. Mention *et al.*, Reactor antineutrino anomaly, *Phys. Rev. D*, 2011 83(7) 073006, http://dx.doi.org/10.1103/PhysRevD.83.073006, publisher: American Physical Society

[205] P. Huber, Reactor antineutrino fluxes – Status and challenges, *Nuclear Physics B*, 2016 908 268, http://dx.doi.org/10.1016/j.nuclphysb.2016.04.012

[206] F. P. An *et al.*, A side-by-side comparison of Daya Bay antineutrino detectors, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators,*

Spectrometers, Detectors and Associated Equipment, 2012 685 78, http://dx.doi.org/10.1016/j.nima.2012.05.030

[207] A. Abusleme *et al.*, The design and sensitivity of JUNO's scintillator radiopurity pre-detector OSIRIS, *Eur. Phys. J. C*, 2021 81(11) 973, http://dx.doi.org/10.1140/epjc/s10052-021-09544-4

[208] C. Genster, Software and hardware development for the next-generation liquid scintillator detectors JUNO and OSIRIS, Ph.D. thesis, Aachen, 2019, medium: online

[209] V. Hejny M. Hartmann, Rootsorter: A new analysis framework for anke, *IKP Annual Report*, 2002

[210] S. Agostinelli *et al.*, Geant4—a simulation toolkit, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2003 506(3) 250, http://dx.doi.org/10.1016/S0168-9002(03)01368-8

[211] M. Agostini *et al.*, The Monte Carlo simulation of the Borexino detector, *Astroparticle Physics*, 2018 97 136, http://dx.doi.org/10.1016/j.astropartphys.2017.10.003

[212] Borexino Collaboration *et al.*, Improved measurement of 8B solar neutrinos with $1.5 kt y$ of Borexino exposure, *Phys. Rev. D*, 2020 101(6) 062001, http://dx.doi.org/10.1103/PhysRevD.101.062001, publisher: American Physical Society

[213] M. Agostini *et al.*, Identification of the cosmogenic ^{11}C background in large volumes of liquid scintillators with Borexino, *Eur. Phys. J. C*, 2021 81(12) 1075, http://dx.doi.org/10.1140/epjc/s10052-021-09799-x

[214] E. G. Adelberger *et al.*, Solar fusion cross sections. II. The pp chain and CNO cycles, *Rev. Mod. Phys.*, 2011 83(1) 195, http://dx.doi.org/10.1103/RevModPhys.83.195, publisher: American Physical Society

[215] J. N. Bahcall, The luminosity constraint on solar neutrino fluxes, *Phys. Rev. C*, 2002 65(2) 025801, http://dx.doi.org/10.1103/PhysRevC.65.025801, publisher: American Physical Society

[216] V. Castellani *et al.*, Solar neutrinos: beyond standard solar models, *Physics Reports*, 1997 281(5) 309, http://dx.doi.org/10.1016/S0370-1573(96)00032-4

[217] A. C. Re, Low energy solar-nu: expected and experimental values for flux, survival probability, and interaction rate in borexino, *Internal Borexino Doc*, 2020

[218] P. C. de Holanda, W. Liao A. Y. Smirnov, Toward precision measurements in solar neutrinos, *Nuclear Physics B*, 2004 702(1) 307, http://dx.doi.org/10.1016/j.nuclphysb.2004.09.027

[219] V. Antonelli *et al.*, Solar Neutrinos, *Advances in High Energy Physics*, 2013 2013 e351926, http://dx.doi.org/10.1155/2013/351926, publisher: Hindawi

[220] N. Vinyoles *et al.*, A New Generation of Standard Solar Models, *ApJ*, 2017 835(2) 202, http://dx.doi.org/10.3847/1538-4357/835/2/202, publisher: American Astronomical Society

[221] I. Esteban *et al.*, Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity, *J. High Energ. Phys.*, 2017 2017(1) 87, http://dx.doi.org/10.1007/JHEP01(2017)087

[222] F. L. Villante *et al.*, A step toward CNO solar neutrino detection in liquid scintillators, *Physics Letters B*, 2011 701(3) 336, http://dx.doi.org/10.1016/j.physletb.2011.05.068

[223] V. Di Marcello *et al.*, Fluid-dynamics and transport of 210Po in the scintillator Borexino detector: A numerical analysis, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2020 964 163801, http://dx.doi.org/10.1016/j.nima.2020.163801

[224] C. C. Miller J. Walker, The Stokes-Einstein law for diffusion in solution, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 1924 106(740) 724, http://dx.doi.org/10.1098/rspa.1924.0100, publisher: Royal Society

[225] W. D. Penny, J. Mattout N. Trujillo-Barreto, CHAPTER 35 - Bayesian model selection and averaging, K. Friston *et al.*, eds., Statistical Parametric Mapping, Academic Press, London, 2007 454–467, http://dx.doi.org/10.1016/B978-012372560-8/50035-8

[226] R. E. Kass A. E. Raftery, Bayes Factors, *Journal of the American Statistical Association*, 1995 90(430) 773, http://dx.doi.org/10.1080/01621459.1995.10476572, publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476572

[227] G. Cowan *et al.*, Asymptotic formulae for likelihood-based tests of new physics, *Eur. Phys. J. C*, 2011 71(2) 1554, http://dx.doi.org/10.1140/epjc/s10052-011-1554-0

[228] C. R. Jenkins J. A. Peacock, The power of Bayesian evidence in astronomy, *Monthly Notices of the Royal Astronomical Society*, 2011 413(4) 2895, http://dx.doi.org/10.1111/j.1365-2966.2011.18361.x

[229] F. Feroz M. P. Hobson, Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses, *Monthly Notices of the Royal Astronomical Society*, 2008 384(2) 449, http://dx.doi.org/10.1111/j.1365-2966.2007.12353.x

[230] F. Feroz, M. P. Hobson M. Bridges, MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics, *Monthly Notices of the Royal Astronomical Society*, 2009 398(4) 1601, http://dx.doi.org/10.1111/j.1365-2966.2009.14548.x

[231] F. Feroz *et al.*, Importance Nested Sampling and the MultiNest Algorithm, *The Open Journal of Astrophysics*, 2019 2 11120, http://dx.doi.org/10.21105/astro.1306.2144, publisher: Maynooth Academic Publishing

[232] J. Skilling, Nested Sampling, *AIP Conference Proceedings*, 2004 735(1) 395, http://dx.doi.org/10.1063/1.1835238, publisher: American Institute of Physics

[233] F. Neitzel, N. Ezhov S. Petrovic, Total Least Squares Spline Approximation, *Mathematics*, 2019 7(5) 462, http://dx.doi.org/10.3390/math7050462, number: 5 Publisher: Multidisciplinary Digital Publishing Institute

[234] P. K. Sinervo, Definition and Treatment of Systematic Uncertainties in High Energy Physics and Astrophysics, *SLAC eConf*, 2003 8

[235] H. Behrens L. Szybisz, On the beta-decay 210Bi(RaE) -> 210Po(RaF), *Nuclear Physics A*, 1974 223(2) 268, http://dx.doi.org/10.1016/0375-9474(74)90414-X

[236] A. G. Carles A. G. Malonda, Precision measurement of the RaE shape factor, *Nuclear Physics A*, 1996 596(1) 83, http://dx.doi.org/10.1016/0375-9474(95)00381-9

[237] H. DANIEL, Shapes of Beta-Ray Spectra, *Rev. Mod. Phys.*, 1968 40(3) 659, http://dx.doi.org/10.1103/RevModPhys.40.659, publisher: American Physical Society

[238] D. Flothmann *et al.*, beta-Spektroskopie mit Halbleiterdetektoren beim Zerfall von32P,49Sc,204Tl und210Bi, *Z. Physik*, 1969 225(2) 164, http://dx.doi.org/10.1007/BF01392517

[239] F. Capozzi *et al.*, Current unknowns in the three-neutrino framework, *Progress in Particle and Nuclear Physics*, 2018 102 48, http://dx.doi.org/10.1016/j.ppnp.2018.05.005

[240] Davide Basilico, Alexandre Göttel Luca Pelicci, Backgrounds estimation for JUNO low and intermediate energy solar neutrino analysis, *JUNO-internal document*, 2021

[241] K. Winger *et al.*, A new compilation of the atmospheric 85krypton inventories from 1945 to 2000 and its evaluation in a global transport model, *Journal of Environmental Radioactivity*, 2005 80(2) 183, http://dx.doi.org/10.1016/j.jenvrad.2004.09.005

[242] Borexino Collaboration *et al.*, CNO and pep neutrino spectroscopy in Borexino: Measurement of the deep-underground production of cosmogenic 11c in an organic liquid scintillator, *Phys. Rev. C*, 2006 74(4) 045805, http://dx.doi.org/10.1103/PhysRevC.74.045805, publisher: American Physical Society

[243] KamLAND Collaboration *et al.*, Production of radioactive isotopes through cosmic muon spallation in KamLAND, *Phys. Rev. C*, 2010 81(2) 025807, http://dx.doi.org/10.1103/PhysRevC.81.025807, publisher: American Physical Society

[244] G. Bellini *et al.*, Cosmogenic Backgrounds in Borexino at 3800 m water-equivalent depth, *J. Cosmol. Astropart. Phys.*, 2013 2013(08) 049, http://dx.doi.org/10.1088/1475-7516/2013/08/049, publisher: IOP Publishing

[245] A. Savitzky M.J.E. Golay, Smoothing and Differentiation of Data by Simplified Least Squares Procedures., *Analytical chemistry*, 1964 36 1627

[246] D. Basilico *et al.*, JUNO sensitivity to intermediate energy solar neutrinos, 7Be, pep, and CNO: results of the two independent analysis performed by Milano and Juelich, *JUNO-internal document*, 2022

[247] Alexandre Göttel *et al.*, JUST: A neutrino fit software for JUNO's solar analysis and sensitivity, *Il Nuovo Cimento C*, 2021 45(1) 1, http://dx.doi.org/10.1393/ncc/i2022-22009-8

[248] S. Baker R. D. Cousins, Clarification of the use of CHI-square and likelihood functions in fits to histograms, *Nuclear Instruments and Methods in Physics Research*, 1984 221(2) 437, http://dx.doi.org/10.1016/0167-5087(84)90016-4

[249] F. Vissani, Luminosity constraint and entangled solar neutrino signals, Solar Neutrinos, WORLD SCIENTIFIC, 2019 121–141, http://dx.doi.org/10.1142/9789811204296_0006

[250] J. Bergström *et al.*, Updated determination of the solar neutrino fluxes from solar neutrino data, *J. High Energ. Phys.*, 2016 2016(3) 132, http://dx.doi.org/10.1007/JHEP03(2016)132

[251] E. H. Bellamy *et al.*, Absolute calibration and monitoring of a spectrometric channel using a photomultiplier, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1994 339(3) 468, http://dx.doi.org/10.1016/0168-9002(94)90183-X

[252] B. K. Lubsandorzhiev Y. E. Vyatchin, Studies of "Kapustinsky's" light pulser timing characteristics, *J. Inst.*, 2006 1(06) T06001, http://dx.doi.org/10.1088/1748-0221/1/06/T06001

[253] J. S. Kapustinsky *et al.*, A fast timing light pulser for scintillation detectors, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1985 241(2) 612, http://dx.doi.org/10.1016/0168-9002(85)90622-9

[254] J.-H. Cheng *et al.*, Determination of the total absorption peak in an electromagnetic calorimeter, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2016 827 165, http://dx.doi.org/10.1016/j.nima.2016.05.010

[255] S. E. Furberg, The Decay Scheme of Zinc-65, *Nature*, 1951 168(4284) 1005, http://dx.doi.org/10.1038/1681005c0, number: 4284 Publisher: Nature Publishing Group

[256] V. Hejny, M. Hartmann A. Mussgiller, RootSorter: A New Analysis Framework for ANKE, 2002

[257] M. Balata *et al.*, The water purification system for the low background counting test facility of the Borexino experiment at Gran Sasso, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1996 370(2) 605, http://dx.doi.org/10.1016/0168-9002(95)00862-4

[258] L. Bieger *et al.*, Potential for a precision measurement of solar pp neutrinos in the Serappis Experiment, *arXiv:2109.10782 [hep-ex, physics:physics]*, 2022 ArXiv: 2109.10782

[259] S. Marcocci, Precision measurement of Solar neutrino fluxes with Borexino and prospects for neutrinoless double beta decay search with 136Xe-loaded liquid scintillators., Ph.D. thesis, INFN Milano, 2018

[260] C. Arpesella *et al.*, Measurements of extremely low radioactivity levels in BOREXINO, *Astroparticle Physics*, 2002 18(1) 1, http://dx.doi.org/10.1016/S0927-6505(01)00179-7

[261] R. A. Jarvis, On the identification of the convex hull of a finite set of points in the plane, *Information Processing Letters*, 1973 2(1) 18, http://dx.doi.org/10.1016/0020-0190(73)90020-3

Eidesstattliche Erklärung

Ich, Alexandre Sébastien Göttel,

erkläre hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt:

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt;
2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;
3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;
4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;
5. Alle wesentlichen Quellen von Unterstützung wurden benannt;
6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;
7. Ein Teil oder Teile dieser Arbeit wurden zuvor veröffentlicht und zwar in:
   - L. Bieger et al., Potential for a precision measurement of solar pp neutrinos in the Serappis Experiment, arXiv:2109.10782 and submitted to Eur. Phys. J. C
   - A. Abusleme et al., The design and sensitivity of JUNO's scintillator radiopurity pre-detector OSIRIS, Eur. Phys. J. C 81, article number: 973 (2021)
   - M. Agostini et al., Experimental evidence of neutrinos produced in the CNO fusion cycle in the Sun, Nature 587 (2020) 577-582
   - M. Agostini et al., Sensitivity to neutrinos from the solar CNO cycle in Borexino, Eur. Phys. J. C 80, article number: 1091 (2020)

Datum

Unterschift

Acknowledgements / Danksagung

Having reached the end of a more than three year journey, I am deeply grateful for having embarked on it, and for all the help I have gotten along the way.

First of course I would like to thank Livia for creating this opportunity and especially for always being open to talk not only about the work itself but also engaging with all kinds of my physics ideas - even when they don't always make sense. I am very grateful for all that I have learned through the different projects, for having experienced working with two completely different collaborations, and though they were cut short by a pandemic I am extremely grateful for all the travels I was a part of, and for the trust that was placed in me. I would like to thank everyone in the neutrino group - I could not have wished for a more dynamic environment and can only regret that we couldn't meet in the office during the past two years.

I'd especially like to thank Christoph for being a great tutor and going above and beyond with the knowledge and coding inspiration, for being an amazing travel companion, from hiking along the chinese wall to hiding from Tycoons, and in general for being such a role model. From the times in the office I would also like to thank my other JUNO office-mates: Philipp, thanks for being the rubber duck when needed and for your patience, and Michaela, thanks for the advice. Speaking of JUNO, thanks to Tobi, it's always a pleasure and I wish that I could have gone with you to install the PMTs. Thanks also to my Borexino office-mates: Mariia, thanks for being so amazing and I'm so glad us two python-users we able to lead the resistance away from the ROOT empire, Sale, thanks for the laughs and the cake(!). Especially thanks to Sindhu for everything we taught each other, and for being a great friend. I'm incredibly grateful that we could work together on the LPoF analysis - which is orders of magnitude better than what it would have been without you.

I can't thank everyone as thoroughly as I would like, but I would also like to thank Ömer for the statistics discussions and entertaining car rides, Giulio for the great advice in Italian food, and not forgetting the people I've had too short contact with, Zara, Yuhang, Yu, Yaping, Davide. Thanks of course to Barbara for all the support and all the organisation in Milano. I would also like to thank the "new generation" who I am sure will take the work to the next level. Thanks Cornelius! I wish I could have installed the ACU but thanks to you I know that it is in capable hands. Thanks Anita for being such a great learner and building JUST with me. I had been wanting to write it for some time and couldn't have wished for a better co-developer. Thanks Luca, I have no doubt you'll make a great representative, thanks Mariam for your inexhaustible enthusiasm and cheerfulness (bonus: in french!), thanks Ricky and Apeksha, thanks Antonia for

suggesting Peggy Gou to me and congratulations - I am writing this hours after your defence.

Furthermore, I'd like to thank Achim for being a great teacher and always being open for discussion, as well as Christopher and the whole group from 3B - I miss the live meetings! Thanks Christian and Jochen for the help with my LED! Thanks also to Michi, not only for being the best leader OSIRIS could ever want but also for always being so supportive.

I'd also like to thank all the people that supported me outside of work, and to all the friends I made during my eight years in Aachen. I can't name everyone but thanks especially to Philipp, Martin, and Anne. I really couldn't have gone this far without you and my life is so much better (and more fun) with you in it.

Of course, thanks to my family for always supporting my scientific endeavours and putting me on the path without which none of this would have been possible. Thanks Maman for all the space news, thanks Papa for all the interesting articles, thanks Thomas for the amazing books, thanks Anne-Sophie for the great support and asking me math questions. Thanks Marlene for the endless amounts of love and support, and for everything else.